National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

MOLECULAR ENZYMOLOGY — CHARACTERIZATION OF TWO MITOCHONDRIAL ENZYMES BY DNA CLONING TECHNIQUES

BY

TIANWEI LIN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

SPRING, 1992

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-73019-6

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: TIANWEI LIN

TITLE OF THESIS: MOLECULAR ENZYMOLOGY--CHARACTERIZATION OF TWO MITOCHONDRIAL ENZYMES BY DNA CLONING TECHNIQUES

DEGREE: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED: 1992

PERMISSION IS HEREBY GRANTED TO THE UNIVERSITY OF ALBERTA LIBRARY TO REPRODUCE SINGLE COPIES OF THIS THESIS AND TO LEND OR SELL SUCH COPIES FOR PRIVATE, SCHOLARLY OR SCIENTIFIC RESEARCH PURPOSE ONLY.

THE AUTHOR RESERVES OTHER PUBLICATION RIGHTS, AND NEITHER THE THESIS NOR EXTENSIVE EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT THE AUTHOR'S WRITTEN PERMISSION.

FLAT 05, 1/F BLK B
JADE COURT
35A BELCHER'ST
HONG KONG

Date: March 27, 1992

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

THE UNDERSIGNED CERTIFY THAT THEY HAVE READ, AND RECOMMED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH FOR ACCEPTANCE, A THESIS ENTITLED *MOLECULAR ENZYMOLOGY — CHARACTERIZATIONS OF TWO MITOCHONDRIAL ENZYMES BY DNA CLONING TECHNIQUES* SUBMITTED BY TIANWEI LIN IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

William A. Bridger

Neil Madsen
Neil B. Madsen

Randy Read
Randy J. Read

Douglas G. Scraba

Joel Weiner
Joel H. Weiner

Choy L. Hew

Date: March 27, 1992

To

My Parents

ABSTRACT

The techniques of DNA cloning were used to characterize two pig heart mitochondrial enzymes, succinyl-CoA synthetase (SCS) and 3-ketoacid:succinyl-CoA transferase (CoA transferase).

A pig heart cDNA library (λgt11) was constructed. It was determined that the library consisted of 1×10^6 independent clones, among which more than 98% were recombinants, and the average size of the inserts was 1.6kb. Furthermore, 4 full length cDNA clones, coding different proteins, were successfully isolated from the library. These data indicated that the library was of very high quality and a good potential source for isolation of other full length clones.

The isolation of two cDNA clones encoding the α subunit of SCS suggested the presence of at least two isozymes of SCS in pig hearts. These two isoforms of the α subunit of SCS were identical in their predicted amino acid sequences except for a stretch of 19 amino acids in the sequence of one protein that was replaced by a stretch of 36 amino acids in the sequence of the other protein.

Evidence for the authenticity of the two isoforms and the likely explanation of their being produced by the mechanism of alternative RNA splicing were obtained by the means of genomic DNA cloning. Both exon sequences were found in the α subunit gene of SCS with the 57 nucleotide exon preceding the 108 nucleotide exon. The intron between the upstream exon and the 57 nucleotide exon was estimated to be approximately 200bp and the intron between the 108 nucleotide exon and the downstream exon was about 1.3kbp. The intron between the two cassettes was not smaller than 5.5kb.

We have isolated a full-length cDNA clone encoding the cytoplasmic precursor to pig heart mitochondrial CoA transferase, a key enzyme for ketone body catabolism. The deduced amino acid sequence, the first reported for this protein, indicates the presence of a 39-residue mitochondrial signal sequence and a 481-residue mature protein of molecular weight 52,197. CoA transferase is known to be susceptible to proteolytic cleavage to produce a nicked but active enzyme. We have identified the site of proteolysis, and analysis of the sequence in its vicinity suggests that the polypeptide may fold into two domains connected by a highly hydrophilic bridge.

Acknowledgements

First, I would like to express my sincere gratitude to Professors Yan Sixu and Huang Houze of Xiamen University, China. Their guidance has set the foundation of my scientific knowledge. Without their help, motivated solely by academic reasons, I would not have had the chance to study in Canada. Their scholarly norm has been the inspiration to me and many others.

I would like to express my warmest appreciation to my supervisor, Dr. William A. Bridger, for his guidance, encouragement, and endless patience to transform me into an independent scientist.

Special thanks to Dr. William T. Wolodko, from whom I acquired a much better understanding of the necessity of rigour in science and the knowledge of Canadian culture, daily life and the language. His patience and willingness for the numerous questions and discussions has been essential and is gratefully appreciated.

I would like to gratefully thank the following people: Mr. Edward R. Brownie for his expertise in the purification of various enzymes, which was absolutely vital to this project; Mrs. Shirley Shechosky for her kindness and her skillful operation of the FPLC™; Mr. Warren Henning, Mr. Perry D'Obrenan and Mrs. Rita Witford for their part in the sequencing experiments of the project; Dr. Kenneth Roy, Mr. Perry D'Obrenan and Mrs. Rita Witford for the synthesis of oligonucleotides; and Mr. Mike Carpenter for the automated protein sequence analysis.

Helpful discussions with fellow graduate students, David Ryan and Darrin Bailey, and postdoctoral fellows, Susan Williams Lees-Miller, Ramanath Majumdar and Guohua Fang, are also greatly appreciated. The author also would like to express appreciation to David Ryan for kindly providing the pig genomic DNA.

I would like to thank the members of my graduate study committee, for their guidance and help. Special thanks goes to Dr. Douglas Scraba, who has provided encouragement and help since the very beginning and throughout the course of my graduate studies.

Financial support provided by the Department of Biochemistry in the form of a Research Assistantship is gratefully acknowledged.

I would like to thank my parents, brothers and sister for their support and encouragement.

I am indebted to many, many of my friends, for their spiritual support through out the course of the project.

Last, but not the least, thank you, Canada.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION 1

CHAPTER II

MATERIAL AND METHODS 11

(I) Materials 11

(II) General methods 12

(1) Gel electrophoresis 12

(a) SDS PAGE 12

(b) TBE PAGE 13

(c) Agarose gel electrophoresis 14

(d) DNA sequencing gel electrophoresis 14

(2) Transfer of proteins or DNA onto membranes 14

(a) Western blots 14

(b) Electrophoretic transfer of DNA 16

(c) Southern blot 16

(3) Screening of λgt11 libraries with antibodies 17

(4) Mini-preparations of plasmid DNA by alkaline lysis method 18

(5) Preparation of single-stranded DNA of M13 phage.............................. 19

(6) Isolation of DNA fragments from agarose gels.................................. 19

(a) The glass powder method.. 19

(b) The DEAE-membrane method... 20

CHAPTER III

GENERATION OF A PIG HEART cDNA LIBRARY...................................... 21

(1) Introduction.. 21

(2) Considerations for building the pig heart cDNA library in λgt11 27

(3) Experimental procedures... 28

(a) Isolation of total RNA and mRNA.. 28

(b) The synthesis of cDNA.. 32

(c) Generation of the EcoRI sites on the ends of the cDNA................ 33

(d) Separation of cDNA from the free linkers and the completion of the library.. 35

(4) Results and Discussion... 36

CHAPTER IV

cDNA CLONING OF THE PRECURSORS OF THE α SUBUNITS OF PIG HEART SUCCINYL-CoA SYNTHETASE.. 42

(1) Introduction 42

(2) Experimental procedures 45

(3) Results and discussion 48

CHAPTER V

GENOMIC DNA CLONING OF THE GENE OF THE α SUBUNIT OF SCS 60

(1) Introduction 60

(2) Experimental procedures 63

(a) Isolation of genomic DNA clones for the α subunit of SCS 63

(b) Oligonucleotide hybridizations 64

(c) Polymerase chain reactions 65

(3) Results 66

(a) Preliminary analysis of the two genomic DNA clones 66

(b) The authenticity of the 108-cassette and the identification of the intron between the 108-cassette exon and the exon downstream 68

(c) The identification of the intron between the upstream sequence and the 57-cassette 73

(d) The intron between 57 and 108-cassette exons 73

(4) Discussion and conclusion 83

CHAPTER VI

SEQUENCE OF A cDNA CLONE ENCODING PIG HEART MITOCHONDRIAL CoA TRANSFERASE 90

(1) Introduction 90

(2) Material and methods 92

(a) cDNA cloning of the CoA transferase 92

(b) Purification of pig heart CoA transferase 92

(c) Protein sequence determination 92

(3) Results and discussion 94

(a) N-Terminal sequence determination 94

(b) cDNA clone isolation and sequencing 94

(c) CoA transferase sequence 95

CONCLUSIONS AND PERSPECTIVE 104

REFERENCES 106

APPENDIX I. Publications 115

APPENDIX II. Some related research 117

List of Tables

Table I Sizes of various portions of pig heart CoA transferase 102

Table II Amino acid composition of CoA transferase...103

List of Figures

Figure 1. Inserts of Clone#6 22

Figure 2. Alignment of the nucleotide sequence of Clone#6 23

Figure 3. Small inserts of the clones isolated from commercial libraries 26

Figure 4. Production of pig heart cDNA library in λgt11 29

Figure 5. Intact total RNA from pig heart 31

Figure 6. Synthesis of cDNA 34

Figure 7. Fractions from FPLC chromatography 37

Figure 8. Inserts of the randomly picked clones 40

Figure 9. Screening pig heart cDNA library with biotinylated probes 44

Figure 10. Comparison of the amino acid sequences of the α subunit of SCS 49

Figure 11. Nucleotide and amino acid sequences of PS57 and PS108 51

Figure 12. The two protein isoforms produced by alternative splicing 52

Figure 13. Identification of the clones coding for α subunit of SCS with 57 cassettes . 53

Figure 14. SDS PAGE analysis of *E.coli* and pig heart SCS 56

Figure 15. Separation of isozymes of pig heart SCS by FPLC............................ 57

Figure 16. Amplification of a genomic DNA fragment 62

Figure 17. Sal I digestion of λE4 and λE6 69

Figure 18. The DNA fragment containing the intron between 108-cassette and its downstream exon 71

Figure 19. The intron between the 108-cassette and the downstream exon 72

Figure 20. The DNA fragment residing the intron between 57-cassette and its downstream exon 74

Figure 21. The intron between the upstream exon and the 57-cassette 75

Figure 22. Restriction enzyme analysis of λE6 76

Figure 23. The organization of two mutually exclusive cassettes 78

Figure 24. Partial sequence of the genomic DNA of α subunit of SCS 79

Figure 25. Cla I digestion of a DNA fragment from λE6 80

Figure 26. DNA fragments that contain the intron before 108-cassette 82

Figure 27. Digestion of λE4 with Sal I and Xho I 84

Figure 28. The sequencing data of the boundary of 108-cassette 86

Figure 29. SDS PAGE analysis of nicked and unnicked preparations of CoA transferase 93

Figure 30. Strategy used for sequence determination of λT6 96

Figure 31. Sequence of λT6 and deduced amino acid sequence of CoA transferase 97

Figure 32. Hydrophilicity plot of the amino acid sequence of CoA transferse............99

Figure 33. Helical wheel plot of CoA transferase residues 271-295......................100

List of Abbreviations

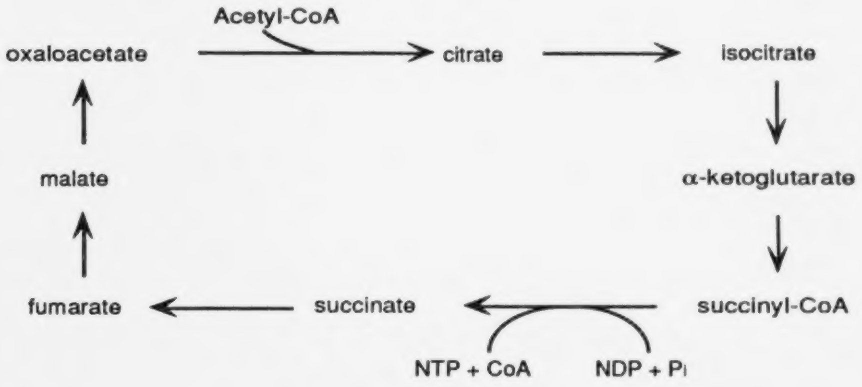
ADP	- adenosine-5'-diphosphate
ATP	- adenosine-5'-triphosphate
BCIP	- 5-bromo-4-chloro-3-indolyl phosphate
bp	- base pairs
BSA	- bovine serum albumin
cDNA	- complementary deoxyribonucleic acid
CoA	- coenzyme A
dATP	- 2'-deoxyadenosine-5'-triphosphate
dCTP	- 2'-deoxycytidine-5'-triphosphate
dGTP	- 2'-deoxyguanosine-5'-triphosphate
DNA	- deoxyribonucleic acid
dNTP	- 2'-deoxynucleotide-5'-triphosphates
DTT	- 1,4-dithiothreitol
dTTP	- 2'-deoxythymidine-5'-triphosphate
dUTP	- 2'-deoxyuridine-5'-triphosphate
EDTA	- ethylenediaminetetraacetate
FPLC™	- fast protein liquid chromatography
GTP	- guanosine-5'-triphosphate
IPTG	- isopropyl-β-thiogalactopyranoside
ITP	- inosine-5'-triphosphate
HPLC	- high performance liquid chromatography
kbp	- kilobase pairs
M.O.I.	- multiplicity of infection
mRNA	- messenger ribonucleic acid
NBT	- 4-nitro blue tetrazolium chloride

NDP	- nucleotide-5'-diphosphate
NTP	- nucleotide-5'-triphosphate
OD	- optical density
PAGE	- polyacrylamide gel electrophoresis
PCR	- polymerase chain reaction
pfu	- plaque forming unit
Pi	- inorganic phosphate
PMSF	- phenylmethylsulfonyl fluoride
RNA	- ribonucleic acid
RNase	- ribonuclease
SCS	- succinyl-CoA synthetase
SDS	- sodium dodecyl sulfate
TCA cycle	- tricarboxylic acid cycle
Tris	- 2-amino-2-(hydroxymethyl)-1,3-propanediol
UV	- ultraviolet
x-gal	- 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER I

INTRODUCTION

The citric acid cycle is the basic energy-supplying metabolic pathway in mammalian systems. As outlined below, it starts with the condensation of acetyl-CoA and oxaloacetate to make citric acid which then has two of its carbons oxidized to CO_2 through the cycle with the regeneration of oxaloacetate (Stryer, 1988).


Acetyl-CoA
oxaloacetate
citrate
isocitrate
α-ketoglutarate
succinyl-CoA
succinate
NTP + CoA
NDP + Pi
fumarate
malate

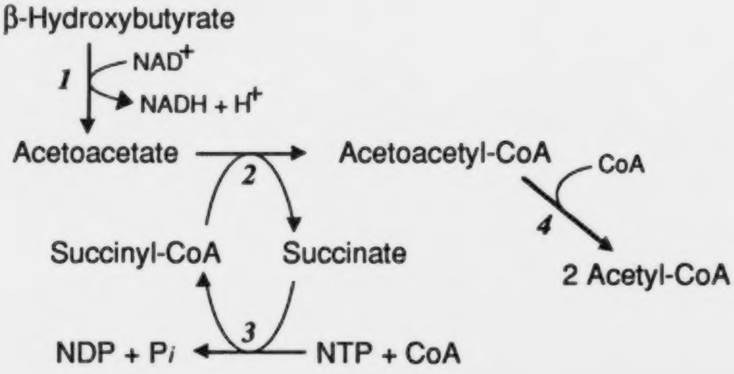
TCA cycle

One major source of acetyl-CoA is from oxidative decarboxylation of the pyruvate produced by glycolysis. However, the constant flow of acetyl-CoA cannot be maintained by glycolysis alone since this process depends on diet and the amount of stored polysaccharides which is limited in all tissues. Therefore most tissues, especially the skeletal muscle and the heart which have major needs for continuous energy supply, require additional suppliers of acetyl-CoA. A variety of energy sources (including fatty acids, lactate and ketone bodies) are used by heart tissues to supply acetyl-CoA, with ketone bodies being one of the principal fuels (Williamson and Krebs, 1961).

The metabolites that are collectively known as ketone bodies include acetoacetate, β-hydroxybutyrate and acetone, with acetoacetate and β-hydroxybutyrate being normal products of hepatic metabolism (β-oxidation) of free fatty acids. Acetoacetate and β-hydroxybutyrate are interconvertible in all mammalian tissues of systems. Although they are produced in the liver, this organ does not have the capacity for their utilization. In other tissues such as the heart, the utilization of acetoacetate is initiated by the action of a mitochondrial enzyme, succinyl-CoA:3-ketoacid CoA-transferase (CoA transferase, EC 2.8.3.5), which catalyzes the following reaction(Jencks, 1973):

$$\text{acetoacetate} + \text{succinyl-CoA} \rightleftharpoons \text{acetoacetyl-CoA} + \text{succinate} \qquad (1)$$

One molecule of acetoacetyl-CoA thus produced can subsequently react with one molecule of Coenzyme A to produce two molecules of acetyl-CoA, and this reaction is catalyzed by acetoacetyl-CoA thiolase (Nosadini *et al.*, 1989). The overall pathway for the conversion of β-hydroxybutyrate and acetoacetate to acetyl-CoA is shown below:



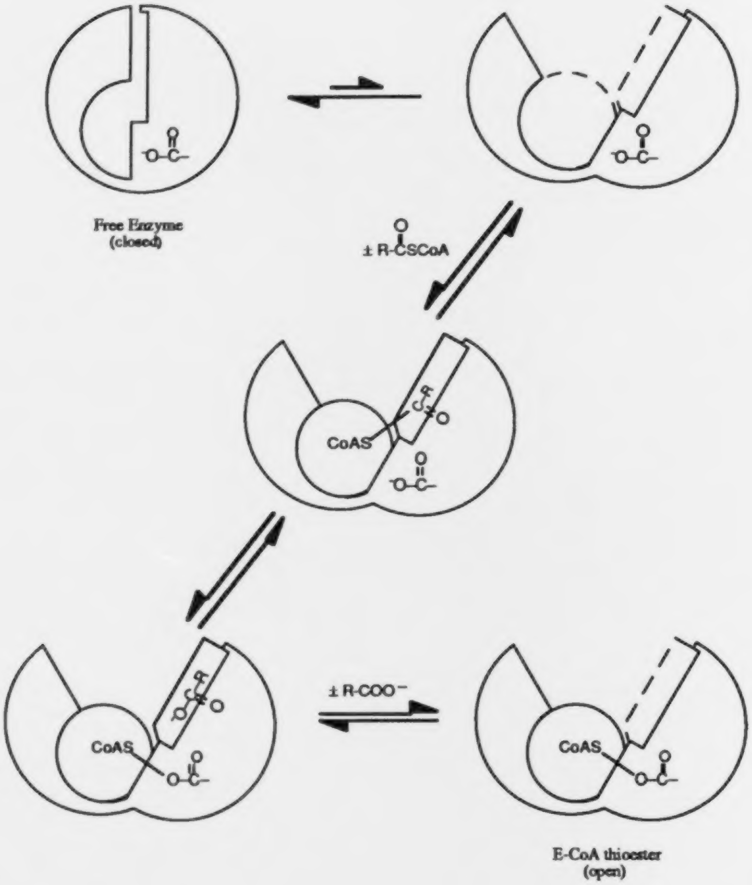
Route for catabolism of β-hydroxybutyrate and acetoacetate. The enzymes involved are: 1, β-hydroxybutyrate dehydrogenase; 2, CoA transferase; 3, succinyl-CoA synthetase; 4, acetoacetyl-CoA thiolase.

Succinyl-CoA:3-ketoacid CoA-transferase (designated in this thesis as CoA transferase) is an enzyme that has escaped comprehensive scrutiny of its structure, with most attention having been focused on its reaction mechanism and kinetics. It was suggested that the enzyme is a homodimer, based upon an estimate of M_r for the subunit molecular weight of 45,600 by sedimentation equilibrium and 52,000 to 63,000 by SDS PAGE (White and Jencks, 1976b). Data on the noncovalent interactions between different moieties of CoA and CoA transferase suggests that each subunit comprises two functional domains (Fierke and Jencks, 1986). It was proposed that the transfer of CoA between succinate and acetoacetate catalyzed by CoA transferase is via two half-reactions involving a covalent enzyme-CoA intermediate (Hersh and Jencks, 1967a):

$$\text{Acetoacetyl-CoA} + \text{enzyme} \rightleftharpoons \text{enzyme-CoA} + \text{acetoacetate} \qquad (2)$$

$$\text{Enzyme-CoA} + \text{succinate} \rightleftharpoons \text{succinyl-CoA} + \text{enzyme} \qquad (3)$$

The reaction follows "ping-pong" kinetics, which is characteristic of enzymatic reactions that proceed through two half-reactions in which the group to be transferred is bound to the enzyme in an intermediate which does not contain the rest of the first substrate. The enzyme-CoA intermediate has been isolated (Hersh and Jencks, 1967b), and the thiolester bond of the intermediate was shown to be formed between the thiol group of the CoA and a γ-carboxyl group of a glutamic acid residue (Solomon and Jencks, 1969). Apparently, at least two conformations could be adopted by the enzyme and an "alligator" model was proposed wherein the conformation of the enzyme is converted from the normal closed form to an open form upon formation of the enzyme-CoA intermediate (White *et al.*, 1976; Pickart and Jencks, 1979). This is shown below. (The diagram is adapted from White *et al.*, 1976)



The interaction between the enzyme and the non-reactive pantetheine moiety of CoA is the major contributor to the rate acceleration in the direction of enzyme-CoA formation forcing the open conformation. The binding energy of carboxylate substrates acts in reducing the energy barrier in the opposite direction to induce a closed conformation (Moore and Jencks, 1982; Fierke and Jencks, 1986).

The continuous production of succinyl-CoA is crucial for the utilization of ketone bodies to maintain the supply of acetyl-CoA in the heart since succinyl-CoA is the donor of the CoA component to acetoacetic acid, as shown in reaction (1). The metabolism of succinyl-CoA involves an enzyme, succinyl-CoA synthetase (SCS) (succinate:CoA ligase, EC 6.2.1.4 [GTP-forming] and EC 6.2.1.5 [ATP-forming]), an enzyme that is well known for its role in the tricarboxylic acid cycle. In order to replenish succinyl-CoA for continued metabolism of acetoacetate, SCS catalyzes the reverse of its TCA cycle reaction:

$$\text{succinate} + \text{CoA} + \text{NTP} \rightleftharpoons \text{succinyl-CoA} + \text{NDP} + P_i \qquad (4)$$

Succinyl-CoA synthetase has been the subject of several reviews (Bridger, 1974; Ottaway *et al.*, 1981; Vogel and Bridger, 1983; Nishimura, 1986). The mechanism of the reaction catalyzed by SCS involves a different covalent intermediate between the enzyme and phosphate. A histidine residue is the site of phosphorylation of the enzyme and the reaction pathway is believed to be as follows:

$$\text{E} + \text{NTP} \rightleftharpoons \text{E-PO}_3 + \text{NDP} \qquad (5)$$

$$\text{E-PO}_3 + \text{succinate} \rightleftharpoons \text{E·succinyl-PO}_3 \qquad (6)$$

$$\text{E·succinyl-PO}_3 + \text{CoA} \rightleftharpoons \text{E}+P_i + \text{succinyl-CoA} \qquad (7)$$

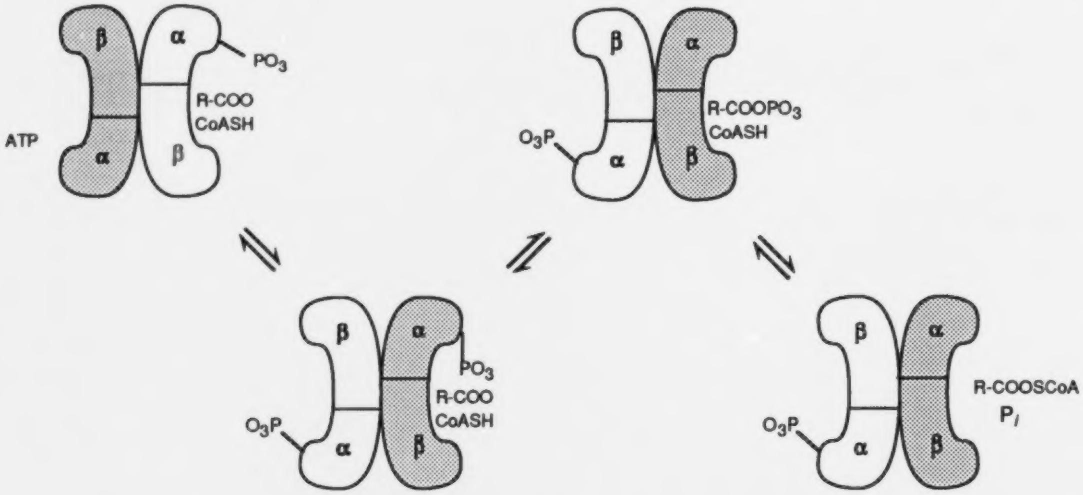
$$\text{Sum:}\quad \text{NTP} + \text{succinate} + \text{CoA} \rightleftharpoons \text{Succinyl-CoA} + \text{NDP} + P_i \qquad (8)$$

E. coli and pig heart SCS are among the best characterized species of this enzyme. The *E. coli* enzyme was shown to be an $\alpha_2\beta_2$ tetramer (Bridger, 1971), and subsequent sequencing of the genes for *E. coli* α and β shows that they encode proteins of molecular weights of 29,644 and 41,390, respectively (Buck *et al.*, 1985). This species of SCS favors ATP as substrate (Gibson *et al.*, 1967), but also reacts with GTP with lower affinity (Murakami *et al.*, 1972). Wang *et al.* (1972) isolated a phosphohistidine-containing peptide and determined its sequence to be:

Met-Gly-His*-Ala-Gly-Ala-Ile-Ile-Ala-Gly-Gly-Lys

This sequence was also found to be encoded in the SucCD clone (Buck, *et al.*, 1985), specifically in the SucD gene, establishing its identity as the gene coding for the α subunit of SCS.

Only one of the two α subunits could be phosphorylated at any given time, giving rise to the possibility of half-site reactivity (Moyer *et al.*, 1967; Moffet *et al.*, 1972). NMR studies (Vogel and Bridger, 1983), hybrid enzyme formation (O'Connor *et al.*, 1985), and single turnover measurements (Wolodko *et al.*, 1983) supported an alternating sites cooperativity model for the action of SCS (Vogel and Bridger, 1982). This is shown below:



The pig heart SCS is believed to be composed of one α and one β subunit with molecular weights estimated to be 34,500 for α and 42,500 for β subunits (Brownie and Bridger, 1972; Murakami and Nishimura, 1974). Except isozymes reportedly isolated from pig heart (Weitzman *et al.*, 1986), until recently all mammalian SCS preparations have been found to be GTP-specific (see below).

The quaternary structures and nucleotide specificities of SCS have been the focus of many studies, and the results have been somewhat confusing. Originally, the tetrameric and ATP-specific *E. coli* SCS was believed to be representative of all bacterial enzymes. However, it was later found that there were varieties of SCS species with different compositions and nucleotide specificities (Weitzman and Jaskowskan-Hodges, 1982; Weitzman, 1987). It was concluded that, with regard to nucleotide specificity, there were four types of bacterial SCS enzymes. The first type was found to have high affinity for GTP but low affinity for ATP. The second kind utilized GTP and ATP equally. The third kind had higher activity for ATP than for GTP, and the fourth kind was specific for ATP only. Also, it was noted that the first, second and third kinds of SCS were all produced by Gram-negative bacteria and they were all tetrameric. However, the fourth kind was only found in Gram-positive bacteria and was dimeric. It was therefore claimed that, in general, there are two kinds of SCS — dimeric and tetrameric — in bacteria, with dimeric enzymes found only in Gram-positive and tetrameric ones only in Gram-negative bacteria (Weitzman, 1987).

Although modern textbooks of biochemistry indicate that mammalian SCS is GTP-specific, recent work in Weitzman's laboratory has demonstrated the presence of ATP-specific SCS in most mammalian tissues (Weitzman *et al.*, 1986; Weitzman, 1987). The different specificities were attributed to two distinct and separable proteins, establishing that there is more than one species of mammalian SCS functioning in various tissues. It is thus possible that the various isoforms of mammalian SCS could be rationalized in terms of the different metabolic roles for this enzyme: participating in the tricarboxylic acid cycle (generating NTP) on the one hand, and participating in ketone body catabolism (consuming NTP) on the other.

Further to this point, it has been noted that rationalizing the kinetic properties of SCS and its metabolic roles could present a dilemma. It has been reported that SCS has a very small K_m value for succinyl-CoA (14μM; Bridger, 1974) compared to that of CoA transferase (4.2 mM); thus SCS possesses a much higher affinity for succinyl-CoA. The TCA cycle *per se* would therefore be favored by the unregulated SCS in pig heart as SCS generates nucleoside triphosphates at the expense of the high energy bond of succinyl-CoA and pushes the cycle forward. However, as the SCS carries the TCA cycle forward, the formation of acetoacetyl-CoA would be diminished as SCS would pre-empt all the available succinyl-CoA and make it unavailable for CoA transferase to generate acetyl-CoA from ketone bodies to supply the TCA cycle. Therefore, McMinn and Ottaway (1979) postulated that uncontrolled TCA cycle activity could impair the pathway for catabolism of ketone bodies. SCS seems to have to play two contradictory roles in the TCA cycle: on one hand it must push reaction (4) forward to produce succinyl-CoA in order to produce acetoacetyl-CoA to supply the starting material to begin the TCA cycle, on the other hand it must keep the reaction (4) going backwards in order to maintain the TCA cycle.

The discovery of two kinds of SCS in mammalian cells led to the hypothesis that the two enzymes could function differently — one to supply succinyl-CoA for the production of acetoacetyl-CoA and the other to participate in the TCA cycle (McClellan and Ottaway, 1980; Ottaway *et al.*, 1981; Jenkins and Weitzman, 1986). Since the ratio of GTP/GDP (at least 100) is much higher than that of ATP/ADP (about 1) in mitochondria, it was suggested that the GTP-specific SCS is responsible for the production of succinyl-CoA and is under the driving force of the higher chemical potential of GTP/GDP (McClellan and Ottaway, 1980). The ATP-specific SCS would catalyze the TCA cycle reaction. Moreover, there was found to be an increase in GTP-specific SCS activity in several tissues of diabetic animals, consistent with the idea that the GTP-specific, not the ATP-specific, SCS could responsible for ketone body metabolism (Jenkins and Weitzman, 1986).

While this hypothesis seems reasonable, it is the author's view that it fails to explain some experimental data. Firstly, the GTP-specific SCS activity in the liver was found to be higher than the ATP-specific activity, despite the fact that there is no ketone body consumption in normal liver cells (Jenkins and Weitzman, 1986; Weitzman, *et al.*, 1986). Secondly, GTP-specific SCS has been found to interact specifically with the α-ketoglutarate dehydrogenase complex which catalyzes the generation of NADH and succinyl-CoA in the preceding step to the TCA cycle (Porpaczy *et al.*, 1983). This would suggest that this particular isoform might actually function preferentially in the TCA cycle and could share metabolites channelling within an enzyme-enzyme complex (Srivastava and Bernhard, 1986). Interestingly, through the investigation of specific interactions between SCS and the α-ketoglutarate dehydrogenase complex, it was found that the K_m values for succinyl-CoA for both SCS and α-ketoglutarate dehydrogenase were decreased, indicating that enzyme-enzyme interactions produced stronger interactions with succinyl-CoA (Porpaczy *et al.*, 1983). Since enzyme-enzyme interaction could change the properties of individual enzymes, and if the transfer of metabolites could proceed through those complexes, the driving force of GTP/GDP could be insignificant inside the mitochondria. The fates of the reaction intermediates might not be influenced by substrates and the end-products in solution, but instead the interactions between enzymes and proteins could be dominant in sharing the metabolic intermediates. It is possible that the compartmentation or localization of a particular species of SCS, rather than its nucleotide specificity, serves as the key element to determine its enzymatic role. However, one cannot rule out the possibility that there is a correlation between the nucleotide specificity and the localization of the enzymes, and further investigations are certainly required.

The possibility of protein-protein interaction between SCS and CoA transferase was examined at an early stage of this project. It was discovered that GTP-specific SCS and CoA transferase could be co-purified and that these enzymes were often found to be contaminants of each other. However, no evidence for specific interactions between these two enzymes was found in co-centrifugation, cross-linking or kinetic studies. It must be pointed out, however, that these studies all used the GTP-specific SCS since no ATP-specific enzyme was then available.

It was obvious that more and better material was required for further studies on the interactions between SCS and CoA transferase. The advances in DNA cloning technology achieved during the 1970s, in particular the advent of DNA sequencing and the development of recombinant DNA methodology (including site-specific mutagenesis and protein engineering) offered powerful new research tools for enzymologists. Using DNA cloning technology could obviously add new dimensions to our understanding of the structure and function of both SCS and CoA transferase.

The first cloning of the SCS gene involved the isolation and sequencing of the SucCD gene of *E. coli* that coded for the β and α sequences of *E. coli* SCS, respectively (Buck *et al.*, 1985; Buck *et al.*, 1986). These genes were subsequently made to be overexpressed in bacteria (Buck and Guest, 1989). Another important development was the cloning and sequencing of the cDNA coding for the cytoplasmic precursor to the α subunit of rat liver SCS (Henning *et al.*, 1988).

As research advanced, it appeared that it would be useful to obtain pig heart clones of the related SCS genes for the following reasons. (a) It had been shown that the *E. coli* SCS follows the alternating sites cooperativity mechanism which requires $\alpha_2\beta_2$ architecture; on the other hand, the possibility of a tetrameric form of pig heart SCS had been ruled out by the experiments of Wolodko *et al.* (1986) and pig heart SCS could therefore not involve subunit cooperativity in its mechanism. The relationship of these different properties of *E. coli* and pig heart SCS had thus attracted attention. Moreover, previous data on SCS had mostly been derived from the studies on *E. coli* and pig heart enzymes, and any further study would naturally be simplified by focusing on the rationalization of the available data with new discoveries about these two enzymes. (b) Since CoA transferase is virtually absent from liver tissues, any interaction that it might have with SCS would most likely be demonstrated in the heart samples.

Based on these facts and arguments, this project was designed at the outset to focus on the cDNA cloning of pig heart SCS and 3-ketoacid: succinyl-CoA CoA-transferase.

CHAPTER II

MATERIALS AND GENERAL METHODS

(I) MATERIALS

Restriction enzymes and the DNA modifying enzymes were purchased from BRL, Boehringer Mannheim or New England Biolabs and were used as recommended by the manufacturers; AMV reverse transcriptase was from Life Science; Sequenase™ and Sequenase™ Version II were from USB; a rat heart cDNA library (in λgt11) and a pig genomic DNA library (in EMBL3) were purchased from Clontech; *Thermus aquaticus* DNA polymerase was from Perkin Elmer Cetus and Amersham; Deoxynucleotides were from Pharmacia; EcoR I linkers (GGAATTCC) were from Pharmacia.

Radioactive chemicals were purchased from Du Pont, Amersham or ICN.

Nitrocellulose membranes were from Schleicher & Schuell and Amersham; Hybond-N nylon membrane was from Amersham; GeneScreen™ hybridization membrane was from Du Pont; Immobilon™ transfer membrane was from Millipore; DEAE-membrane was from Schleicher & Schuell.

Buffer and other reagent chemicals were purchased from Sigma, BDH, Bio-Rad, BRL or ICN and were the best quality available.

E.coli strain HB101 is *hsdS20, supE44, ara14,galK2,lacY1,proA2,rspL20, xyl-5, mtl-1, recA13, mcrB, mcrA,mrr;* JM103 is *endA1, hsdR, supE, sbcBC, thi-1, strA, Δ(lac-pro),[F' traD36, lacI*^{q}*ZΔM15, proAB];* JM109 is *recA, endA1, gyrA96, thi-1, hsdR17,* (r_{k-}, m_{k+})*, supE44, relA1, Δ(lac-proAB), [F' traD36, proAB, lacI*^{q}*ZΔM15];* NM538 has been documented (Frischauf *et al.*, 1983); Y1088 is *Δ(lac)U169, supE, supF, hsdR,* *mcrA, metB, trpR, tonA21, proC::Tn5 (kan^r) [pMC9];* Y1090 is *Δ(lac)U169, Δ(lon), araD139, strA, supF,mcrA, trpC22::Tn10, [pMC9].*

DNA plasmids pUC18 and pUC19 are described by Yanisch-Perron *et al.* (1985); M13 phages M13mp18 and M13mp19 are described by Messing (1983) and Yanisch-Perron *et al.* (1985).

(II) GENERAL METHODS

(1) Gel electrophoresis

(a) SDS PAGE

SDS PAGE is a well established method (Laemmli, 1970). Described here is the specific protocol used throughout this project.

For a final volume of 100 ml of 12% separating gel solution: 40 ml 30% acrylamide and 0.8% N, N'-methylenebisacrylamide (33.3 ml for 10% and 26.7 ml for 8% separating gel solutions), 20 ml 2 M Tris-HCl (pH 8.8), 0.4 ml 25% SDS, 40 μl TEMED and 0.3 ml 10% ammonium persulfate for the polymerization.

For the stacking gel: 6 ml 0.5 M Tris-HCl (pH6.8), 8.3 ml of 30% acrylamide and 0.8% N,N'-methylenebisacrylamide, 0.2 ml 25% SDS, 20 μl TEMED, 12.5 ml 60% sucrose and the final volume was adjusted to 50 ml by H_2O. The polymerization was initiated by adding 0.15 ml 10% ammonium persulfate to the deaerated solution.

Five times (5X) concentrated electrophoresis buffer was made by dissolving 2880 g of glycine and 600 g of Tris base in H_2O to the final volume of 20 L; this could be diluted 5 times and supplemented with 0.1% SDS to make the working solution.

The protein samples were prepared in 50 mM Tris-HCl (pH6.8), 2.5% SDS, 1% β-mercaptoethanol, 6% glycerol and 0.1% bromophenol blue and boiled for 5 minutes before application to the gel. Visualization of the proteins in the gel was achieved by staining the gel in 0.5% Coomassie blue, 50% methanol and 10% acetic acid, and then destaining the gel to reduce the background with 35% methanol and 10% acetic acid.

Gradient SDS PAGE was found to give better resolution as it could yield sharper protein bands and was also used routinely. The gradient separating gel ranging from 7 to 15% acrylamide was generated by using gradient formers. In the outlet chamber of the gradient former, 15% gel solution was prepared by mixing 10 ml 30% acrylamide and 0.8% N,N'-methylenebisacrylamide, 4 ml 2 M Tris-HCl (pH8.8), 80 µl 25% SDS, 8 µl TEMED for a final volume of 20 ml adjusted with 60% sucrose. 7% gel solution was prepared in another chamber from 4.7 ml 30% acrylamide and 0.8% N, N'-methylene-bisacrylamide, 4 ml 2M Tris-HCl pH8.8, 80 µl 25% SDS, 8 µl TEMED to a final volume of 20 ml adjusted with H_2O. Polymerization were initiated by adding 60 µl of 10% ammonium persulfate to the solutions in each chamber.

(b) Tris-Borate-EDTA (TBE) PAGE

TBE PAGE was used for separation and analysis of the small DNA fragments sized from tens to hundreds of base pairs. A 5% gel solution was made from 16.6 ml 30% acrylamide and 1% N,N'-methylene-bisacrylamide, 20 ml 5X TBE and adjusted to a final volume of 100 ml with H_2O. Polymerization was initiated with 1.25 ml 10% ammonium persulfate and 130 µl TEMED. 5xTBE was made from 54 g of Tris base, 27.5 g boric acid and 20 ml of 0.5M EDTA (pH8.0) to a final volume of 1 L, with the adjustment of the pH to 8.3 with boric acid.

(c) Agarose gel electrophoresis

Non-denaturing agarose gels buffered by Tris-acetate-EDTA (TAE), alkaline agarose gels, and formaldehyde agarose gels were prepared and used in electrophoresis according to established methods (Berger and Kimmel, 1987, pp75f) without modification.

(d) DNA sequencing gel electrophoresis

Ultra-thin gel systems (Ansorge and Barker, 1984) with field gradients (Olsson *et al.*, 1984; Ansorge and Labeit, 1984) were used for all DNA sequencing experiments. Unlike the buffer-gradient gels (Hong, 1987), this system improved resolution and required only one kind of gel solution. Once it was formed, the field gradient did not diminish during long period of experiments when required; this allowed one to obtain sequencing information further away from the primers (sequences beyond 1kb from the primer and 600 nucleotides could be analyzed in a gel 65cm in length without double loading).

It was the author's experience that the sliding technique of casting the ultra-thin gels (Ansorge and Maeyer, 1980; Garoff and Ansorge, 1981), which requires the moving of one glass plate over the other while pouring the gel solution instead of casting the gel solution into glass sandwich for the gel of normal thickness, was not required to cast the wedge-shaped ultra-thin gels, since the ends of the wedge gels were thick enough to allow the direct casting of gel solutions into the gel moulds.

(2) Transfer of proteins or DNA onto membranes

(a) Western blots

After protein samples were separated by SDS-PAGE, the gel was equilibrated in transfer buffer (made from 12.1 g of Tris base, 57.6 g glycine and 800ml methanol to the final volume of 4 L) for 30 minutes. The equilibrated gel was assembled into the blotting apparatus (Bio-Rad) with a nitrocellulose filter in close contact. The transfer was performed with the gel facing the negative electrode and the nitrocellulose paper facing the positive electrode. Transfer was accomplished by applying 25 volts overnight followed by 65 volts for one hour.

After the transfer, the nitrocellulose filter was incubated in 5% (w/v) skim milk, 0.01% (v/v) Antiform and 0.02% sodium azide for about 1 hour in room temperature before being rinsed thoroughly in water and air-dried.

The gel, after the transfer, could be stained by Coomassie blue to monitor the efficiency of the transfer. It was our experience that the transfer was usually less than 50% complete in each experiment, and each gel could be used for up to 3 times in transfer experiments.

The protein blotted onto the nitrocellulose filter was detected with specific antibodies combined with a second antibody conjugated with horse-radish peroxidase (or alkaline phosphatase) in a colour development experiment. The nitrocellulose filter blotted with the particular protein was wetted in 50mM Tris-HCl (pH7.5) and 150mM KCl and then incubated with the specific rabbit antibody which had been diluted in the same solution (with sodium azide) overnight at room temperature. The nitrocellulose filter was washed twice in large volumes of 50mM Tris-HCl (pH7.5), 150mM KCl and 0.05% anionic detergent (Tergitol NP-40) and twice in 50mM Tris-HCl (pH7.5) and 150mM KCl, before incubation with goat anti-rabbit IgG conjugated with horse radish peroxidase with proper dilution in 50mM Tris-HCl (pH7.5) and 150mM KCl supplemented with 3% BSA for 2 to 3 hours in 37°C. The nitrocellulose filter was then washed twice in 50mM Tris-HCl (pH7.5) and 150mM KCl, twice in the same solution plus 0.05% Tergitol NP-40, before the colour development in a solution made from 30 mg of 4-chloronaphthol, 10ml methanol, 50 ml water and 30 µl 30% H_2O_2 for about half hour. The reaction was stopped by rinsing the filter in water. The filter could be air-dried for storage.

(b) Electrophoretic transfer of DNA

The electrophoretic transfer of DNA was performed on agarose gels buffered with TAE (10mM Tris-Acetic acid (pH7.8), 5mM sodium acetate and 0.5mM EDTA).

The DNA in the agarose gel was denatured by 0.25 M NaOH for 15 minutes then neutralized by washing twice with large volumes of 4xTAE, 10 minutes each time, before being equilibrated with TAE. The gel thus treated could be assembled into the blotting apparatus with a nylon-based transfer membrane and the transfer was performed with pre-chilled TAE at 80 volts for about 1 hour.

After the experiment, the DNA could be fixed onto the membrane by either UV cross-linking or baking at 80°C.

(c) Southern blot

The electrophoretic transfer of DNA, although generally efficient, was not suitable for yielding blots on nitrocellulose filters since nitrocellulose filters only bound DNA in high salt environments in which excess heat would be generated by the electrophoretic process. Therefore, the traditional Southern blot which did not required high salt environment was used when nylon membranes were not desired, as in the case of probing the blots by biotin-streptavidin system which normally produced higher background in the blots on nylon membranes.

The gel to be used in the transfer was first stained with ethidium bromide, photographed and exposed to UV radiation for a short period of time to nick the DNA. It was then soaked in 1.5 M NaCl and 0.5 M NaOH for 1 hour and neutralized in 1 M Tris-

HCl (pH8.0) and 1.5 M NaCl for 1 hour. The gel thus treated was placed on a wet piece of 3MM paper which was immersed in 10xSSC by two edges and a wet piece of nitrocellulose paper of proper size was placed onto the gel. The nitrocellulose paper was covered by another piece of 3MM paper which in turn was covered by a stack of paper towels with a weight. The DNA was transferred onto the nitrocellulose by the capillary effect produced by the paper towels. The transfer was normally allowed to proceed overnight.

(3) Screening of λgt11 libraries with antibodies

A single colony of Y1090 *E. coli* was used to innoculate 20 ml of LB medium, supplemented with 50 μg/ml of ampicillin and 0.2% maltose. The bacteria were allowed to grow at 37°C overnight with shaking. The bacteria (in stationary phase) were harvested and resuspended in 10 ml of 10mM $MgSO_4$.

0.2 ml of the bacterial suspension was transfected with 1.5×10^4 pfu of λgt11 phages and was incubated at 37°C for 20 minutes before being mixed with 3 ml of 0.7% agar prepared in LB medium and spreaded onto a LB/Ampicillin plate 85mm in diameter. The phages were allowed to replicate at 40°C for about 4 hours before the agar was overlayed by a piece of nitrocellulose paper saturated by 10mM IPTG. After 3 more hours of incubation at 37°C, the filter was marked, removed and incubated with 1% calf hemoglobin prepared in 50mM Tris-HCl (pH7.5) and 150mM KCl in room temperature for 1 hour and then incubated with the specific rabbit antibodies overnight at room temperature. The discs were washed twice in large volumes of 50mM Tris-HCl (pH7.5), 150mM KCl and 0.05% anionic detergent (Tergitol NP-40) and twice in 50mM Tris-HCl (pH7.5) and 150mM KCl, and then incubated with goat anti-rabbit IgG conjugated with horse-raddish peroxidase (from Bio-Rad, diluted 2000 times in 50mM Tris-HCl (pH7.5) and 150mM KCl with 3% BSA) for 2 to 3 hours in 37°C. The nitrocellulose discs were washed twice in 50mM Tris-HCl (pH7.5) and 150mM KCl, twice in the same solution plus 0.05% Tergitol NP-40, before colour development in a solution made from 30 mg of 4-chloronaphthol, 10 ml methanol, 50 ml water and 30 µl 30% H_2O_2 for about half hour. The reaction was stopped by rinsing the filter in water. The filter could be air-dried for storage.

(4) Mini-preparations of plasmid DNA by alkaline lysis method

Two ml of LB medium supplemented with the appropriate antibiotic were inoculated with a single bacterial colony and shaken vigorously at 37°C overnight. About 1.5 ml of the overnight culture was transferred into a Microfuge tube and centrifuged for 1 minute to collect the bacterial cells. The pellet of the bacterial cells was resuspended in 100 µl of 50mM glucose, 10mM EDTA, 25mM Tris-HCl (pH8) and 2 mg/ml lysozyme. It was incubated at room temperature for 5 minutes and was then gently mixed with 200 µl of freshly prepared 0.2 M NaOH and 1% SDS and left at room temperature for 5 minutes. The solution was neutralized by 150 µl of sodium acetate (pH 4.8) for 5 minutes and the supernatant was transferred to a fresh tube after separation from the precipitate by centrifugation. After being extracted with an equal volume of phenol and chloroform 1:1/v:v), the plasmid DNA was precipitated from the supernatant fraction by adding two volumes of 95% ethanol. The pellet was collected by centrifugation and resuspended in TE buffer, then treated with RNase A at a concentration of 20 µg/ml at 37°C for an hour. The plasmid solution was then again extracted with an equal volume of phenol and chloroform, then chloroform before precipitation with one-tenth volume of 3 M sodium acetate and two volumes of 95% ethanol and chilling to -20°C. The pellet was collected by centrifugation and was then rinsed with 70% ethanol at -20°C and dried before resuspension in TE as working solution.

The plasmid thus prepared was found to be suitable for double-stranded DNA sequencing.

(5) Preparation of single-stranded DNA of M13 phage

A single plaque of M13 phage and 15 µl of the overnight culture of JM103 were added to 1.5 ml of the 2xYT culture medium and shaken vigorously at 37°C for 5 to 6 hours. The culture was transferred to a Microfuge tube and centrifuged for 1 minute before the supernatant was transferred to another fresh tube. An aliquot of 50 µl of the supernatant was removed and stored as a phage stock at 4°C. 200 µl of a solution of 2.5 M NaCl and 20% polyethyleneglycol 8000 was added to the remaining phage suspension and it was incubated at room temperature for 15 minutes. The phage pellet was collected by centrifugation for 5 minutes and resuspended in 100 µl TE before extraction with a 50:50 (v/v) mixture of phenol and chloroform then by chloroform. The DNA was precipitated by adding 10 µl 3 M sodium acetate and 250 µl 95% ethanol. The DNA pellet, after a centrifugation was separated from the aqueous phase and was rinsed with 0.5 ml 70% ethanol chilled at -20°C and dried. A yield of 1 to 5 µg of DNA could normally be achieved by this procedure and the DNA was suitable for DNA sequencing and oligonucleotide-directed mutagenesis.

(6) Isolation of DNA fragments from agarose gels

(a) The glass powder method

The glass powder suspension and sodium iodide solution were prepared by established methods (Davis *et al.*, 1986).

DNA fragments were separated by electrophoresis in agarose gel buffered by TAE and a piece of the gel containing the target fragment was removed. The gel stub was melted in 2.5 volumes of a saturated sodium iodide solution at about 50°C and the DNA was allowed to bind to the glass powder by occasionally shaking at room temperature for 5 to 15 minutes. The bound DNA fragment was spun down with the glass powder and separated from the liquid phase. The glass powder was forcefully resuspended twice in 0.5 ml of 20mM Tris-HCl (pH7.4), 1mM EDTA, 0.1 M NaCl and 50% ethanol chilled at -20°C after being centrifuged as a pellet. The DNA fragment was then eluted twice by suspension of the glass powder in TE buffer, heating at 50°C and centrifugation to pellet the glass powder.

(b) The DEAE-membrane method

The DEAE-membrane was pretreated by first washing in 10mM EDTA (pH7.5) for 10 minutes and then in 0.5 M NaOH for 5 minutes before rinsing in water. The membrane thus treated could be stored at 4°C in water for several weeks.

DNA fragments were separated by electrophoresis in agarose gel buffered by TAE and visualized by long-wave length UV (366 nm) after being stained with ethidium bromide. An incision was made in the gel just in the front of the target DNA band and a piece of DEAE-membrane trimmed to proper size was inserted. The gel was placed back in the electrophoresis apparatus and the electrophoresis was carried out further until the DNA was completely bound to the membrane (usually about 5 minutes as monitored by a hand-held UV light). The membrane was removed from the gel and rinsed with 20mM Tris-HCl (pH8.0), 0.1mM EDTA and 0.15 M NaCl and the DNA was eluted from the membrane in 20mM Tris-HCl (pH8.0), 0.1mM EDTA and 1 M NaCl in a Microfuge tube incubated at 55 to 68°C for about an hour with occasional shaking. The eluate could be extracted with n-butanol saturated by water to removed ethidium bromide and precipitated by ethanol or reduced the volume and the salt concentration by spin-dialysis.

CHAPTER III

GENERATION OF A PIG HEART cDNA LIBRARY

(1) INTRODUCTION

The search for the cDNA clone coding for CoA transferase was the first project involving the tools of molecular biology in which the author was engaged. Due to the limited research on this particular enzyme, there was no amino acid sequence information available for the design of a probe to clone the CoA transferase gene. However, the successful purification of CoA transferase by Mr. Edward Brownie in this laboratory made it feasible to obtain the information required for the design of oligonucleotide probes (Chapter VI), and to produce antibodies that could be used for probing an appropriate cDNA expression library.

The first probes available for isolating the cDNA of CoA transferase were polyclonal antibodies produced by injecting the rabbits with the purified proteins. Screening rat heart and liver cDNA libraries in λgt11 (kindly provided by M. Mueckler, Massachusetts Institute of Technology, Cambridge, MA), which were the only ones available to us at that time, with the antibodies yielded a few clones that showed positive signals. Among these clones, one designated Clone #6 from rat liver cDNA library was found to give the strongest signal and had the biggest insert about 900 bp (Figure 1). It was therefore selected for sequencing. However, analysis of this sequence using data bases showed that it was 94.7% identical to the nucleotide sequence of the mRNA for the catalytic subunit of the mouse cAMP-dependent protein kinase (Figure 2), and the longest reading frame of Clone #6 corresponded to the protein sequence of the catalytic subunit of

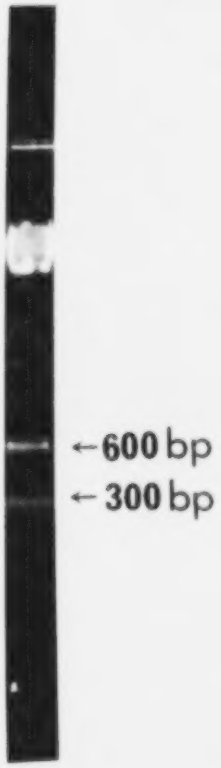


Figure 1. Inserts of Clone#6 The λ DNA of Clone#6 was digested with EcoR I and separated on a 0.8% agarose gel. The insert of the Clone#6 was splited into two fragments of 600bp and 300bp, respectively, by EcoR I.

```
   1'                            GAATTCTCCTTCAAGGACAACTCAAACTTGTAC
                                 ::::::::::::::::::::::::::: :::::
 481" GTCAACTTCCCGTTCCTGGTCAAACTTGAATTCTCCTTCAAGGACAACTCAAACCTGTAC


  34' ATGGTCATGGAGTACGTGCCTGGTGGGGAGATGTTCTCCCACTTACGGCGGATTGGGAGG
      :::::::::::::: ::  ::::::: ::::::::::::::: ::::::::::::: :::
 541" ATGGTCATGGAGTATGTAGCTGGTGGCGAGATGTTCTCCCACCTACGGCGGATTGGAAGG

  94' TTCAGTGAGCCCCACGCCCGTTTCTACGCGGCGCAGATCGTCTTGACTTTTGAGTATCTG
      ::::: :::::::: ::::::::::::::::::::::::::: :::: ::::::::::::
 601" TTCAGCGAGCCCCATGCCCGTTTCTACGCGGCGCAGATCGTCCTGACCTTTGAGTATCTG

 154' CACTCCCTGGACCTCATCTACCGGGACCTGAAGCCCGAGAACCTTCTCATCGACCAGCAG
      ::::::::::::::::::::::::::::::::::::::::: ::::::::::::::::::
 661" CACTCCCTGGACCTCATCTACCGGGACCTGAAGCCCGAGAATCTTCTCATCGACCAGCAG

 214' GGCTATATTCAGGTGACAGACTTCGGTTTTGCCAAGCGTGTGAAAGGCCGAACTTGGACC
      :::::::::::::::::::::::::::::::::::::::::::::::::: :::::::::
 721" GGCTATATTCAGGTGACAGACTTCGGTTTTGCCAAGCGTGTGAAAGGCCGTACTTGGACC

 274' TTGTGTGGGACCCCCGAGTACTTGGCCCCCGAGATTATCCTGAGCAAAGGCTACAACAAA
      :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
 781" TTGTGTGGGACCCCTGAGTACTTGGCCCCCGAGATTATCCTGAGCAAAGGCTACAACAAG

 334' GCTGTGGACTGGTGGGCTCTCGGGGTCCTCATCTACGAGATGGCTGCTGGTTACCCGCCC
      ::::::::::::::::::::::: :::::::::::::::::::::::::::::::: :::
 841" GCTGTGGACTGGTGGGCTCTCGGAGTCCTCATCTACGAGATGGCTGCTGGTTACCCACCC

 394' TTCTTCGCTGACCAGCCTATCCAGATCTATGAGAAAATCGTCTCTGGGAAGGTGCGGTTC
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 901" TTCTTCGCTGACCAGCCTATCCAGATCTATGAGAAAATCGTCTCTGGGAAGGTGCGGTTC

 454' CCATCCCACTTCAGCTCTGACTTGAAGGACCTGCTTCGGAACCTTCTGCAGGTGGATCTC
      ::::::::::::::::::::::::::::::::::: :::::::::::::: ::::::::
 961" CCATCCCACTTCAGCTCTGACTTGAAGGACCTGCTGCGGAACCTTCTGCAAGTGGATCTA

 514' ACCAAGCGCTTTGGGAATCTCAAGAACGGGGTCAATGACATCAAGAACCACAAGTGGTTT
      :::::::::::::: :: :::::: :::::::::::::::::::::::::::::::::::
1021" ACCAAGCGCTTTGGAAACCTCAAGGACGGGGTCAATGACATCAAGAACCACAAGTGGTTT

 574' GCCACAACTGACTGGATCGCCATCTATCAGAGAAAGGTGGAAGCTCCCTTCATACCAAAG
      ::::: ::::::::::: ::::::::::::::::::::::::::::::::::::::::::
1081" GCCACGACTGACTGGATTGCCATCTATCAGAGAAAGGTGGAAGCTCCCTTCATACCAAAG

 634' TTTAAAGGCCCTGGGGACACGAGTAACTTTGACGACTATGAGGAGGAAGAGATCCGGGTC
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1141" TTTAAAGGCCCTGGGGACACGAGTAACTTTGACGACTATGAGGAGGAAGAGATCCGGGTC

 694' TCCATCAATGAGAAATGTGGCAAGGAGTTTACTGAGTTTTAGGGGTGTGCTTGTGCCCCA
      :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
1201" TCCATCAATGAGAAGTGTGGCAAGGAGTTTACTGAGTTTTAGGGGTGTGCTTGTGCCCCT
```

```
  754' TGGGTTTTCTTTTCTTTTTTTTTTTTTTTTCTTTTTCTTTCTATTTTTTTTTCTGGTTGG
       ::::          :: : :::  ::::::::::::::::::: ::::::::: ::::::
 1261" TGGG----------TTCTCTTTCATTTTTTCTTTTTCTTTCTA-TTTTTTTTCCGGTTGG
                                                 |

  814' GGGTGGGGGGGTTGGG
       ::::::: :::::::
 1310" GGGTGGGAGGGTTGGATCGGAACAGCCAGAGGGCCCTAGAGTTCCATGCATCTAATTTAA
```

Figure 2. Alignment of the nucleotide sequence of Clone#6 with that of mouse cAMP-dependent protein kinase catalytic subunit The alignment was done by program ALIGN of BIONETTM. The upper sequence is the sequence of Clone#6 and the lower sequence is that coding for mouse cAMP-dependent protein kinase catalytic subunit. There is 94.7% identiy with 828 nucleotides overlap in the alignment.

mouse cAMP-dependent protein kinase. It was obvious that Clone #6 most likely encoded the catalytic subunit of rat cAMP-dependent protein kinase rather than CoA transferase and was positive in the screening due to the cross-reactivity of the antibodies. It was not studied further. Because of the fact that the liver has low CoA transferase activity, no further sequencing experiments were carried out on other clones from this library.

As the research progressed, two fragments of the CoA transferase arising from a possible protease degradation during purification were separated (Chapter VI) and subjected to protein sequencing. Based on the first available protein sequence of the enzyme, which corresponded to amino acid residues 12 to 17 of the mature enzyme (Chapter IV), an oligonucleotide probe was designed and synthesized. This represented the anti-mRNA sequences and was a 17mer with degeneracy of 128 species, as shown below:

5'CC(AG)TT(GATC)GG(GATC)AT(AG)TC(CT)TT

No desired clones were obtained from a commercial rat heart cDNA library (Clontech) by hybridization experiments with this probe, despite the fact that the conditions of the hybridization experiments were at relatively high stringency (Suggs *et al.*, 1981).

Several lessons could be learned through this "trial and all error" process even though the failure could be conveniently be attributed to the lack of luck. First, every effort should be made to ensure the best quality of the probes and, whenever possible, the use of multiple probes should be used to verify the authenticity of the positive signals. False positives could be commonplace phenomena when screening a library, since there are many genes in each library and some of them, related or not, might produce gene products that are immunologically cross-reactive. Also, polyclonal antiserum could contain antibodies specific for contaminants in the immunizing antigen. It would be prudent to use the sera raised from at least two animals and check the cross-reactivity of the antibodies, even though it might not reveal minor contaminations as had been experienced by the author. The techniques of affinity purification of the antibodies (Smith and Fisher, 1984) could be applied when the antibodies were found to be contaminated. As for oligonucleotide probes, efforts should be spent to derive them from the most reliable protein sequences. It is, of course, advisable to avoid amino acid residues such as serine, leucine and arginine which have a high degeneracy of codon usage when designing a relatively short probe containing all the possible species. If those kind of residues must be be included, the design principle of the so-called "guessmers" (Lathe, 1985) should be applied. In any event, more than one oligonucleotide should be used. A combination of using antibodies first and then oligonucleotide probes for the positive identification of clones is also recommended.

Second, the screening conditions should be carefully investigated. There are no set rules dictating the best hybridization conditions for a particular oligonucleotide probe. Formulas for deriving the hybridization conditions are empirical and can only serve as guidelines for the starting conditions since two oligonucleotides would likely behave differently even if they had the same nucleotide content but different sequences.

Third, I believe that a good DNA library is the single most important factor in the whole process. Likelihood of success in cloning the right gene could be greatly improved if one is using a good library even under unfavorable conditions such as with cross-reacting probes. We have found that some of the commercial libraries are dominated by very small inserts. In our experience with commercial libraries, many positive clones were isolated and were only found to have inserts of about 100bp (Figure 3). These small clones were probably the products of a cDNA library of careless construction (for example, use of degraded mRNA or sub-optimal conditions for cDNA synthesis) and extensive amplification. They would also have a better chance to show a positive signal since there would be more copies in each plaque.

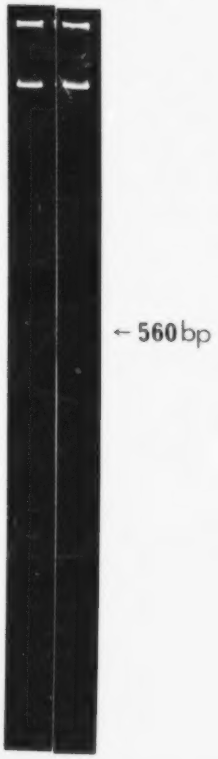


Figure 3. Small inserts of the clones isolated from commercial libraries The λ DNA was digested by EcoR I and separated in a 5% PAGE buffered by TBE. The sizes of the inserts of the two clones were estimated about 100 to 150bp. The library was rat heart cDNA library from Clontech.

Based on these ideas, efforts were expended toward the optimization of the experimental procedures involved in obtaining a good cDNA library. This is the subject of this Chapter.

(2) Considerations for building the pig heart cDNA library in λgt11

The first factor to be considered should be the source of the mRNA. In this case, it was straightforward since the heart is a rich source of both SCS and CoA transferase and most of the characterization of these enzymes has been done using pig heart as the source of material. So porcine heart was the natural choice as the starting organ for the isolation of mRNA. Comparing several preparations, we found that fetal pig hearts had higher RNA content than those of mature animals, consistent with a high rate of growth and protein synthesis.

The second factor to be considered was the type of vector to be used and the methodology for cDNA synthesis. While the classical method of synthesizing cDNA involving S1 nuclease digestion was considered to be inefficient (Rougeon and Mach, 1976; Efstratiadis *et al.*, 1976), the method of Okayama and Berg (1982) was elaborate and required the use of plasmid vectors; this would involve colony hybridization for screening and would be more tedious than plaque hybridization. This method was also ruled out despite its inherent advantage of cloning full length cDNA (at the expense of total clones). It was our belief that the method of cDNA synthesis employed by Gubler and Hoffman (1983) was the best available due to its simplicity, reasonable chance of cloning full length cDNA and the freedom of choice of vectors. The λgt11 system (Huynh *et al.*, 1985) was chosen as the vector for the library. This vector has been proven to be efficient, reliable, and could offer the versatility of being screened with both antibody and oligonucleotide probes.

Figure 4 shows a schematic representation of the generation of the pig heart cDNA library, and the experimental procedures used are summarized below.

(3) EXPERIMENTAL PROCEDURES

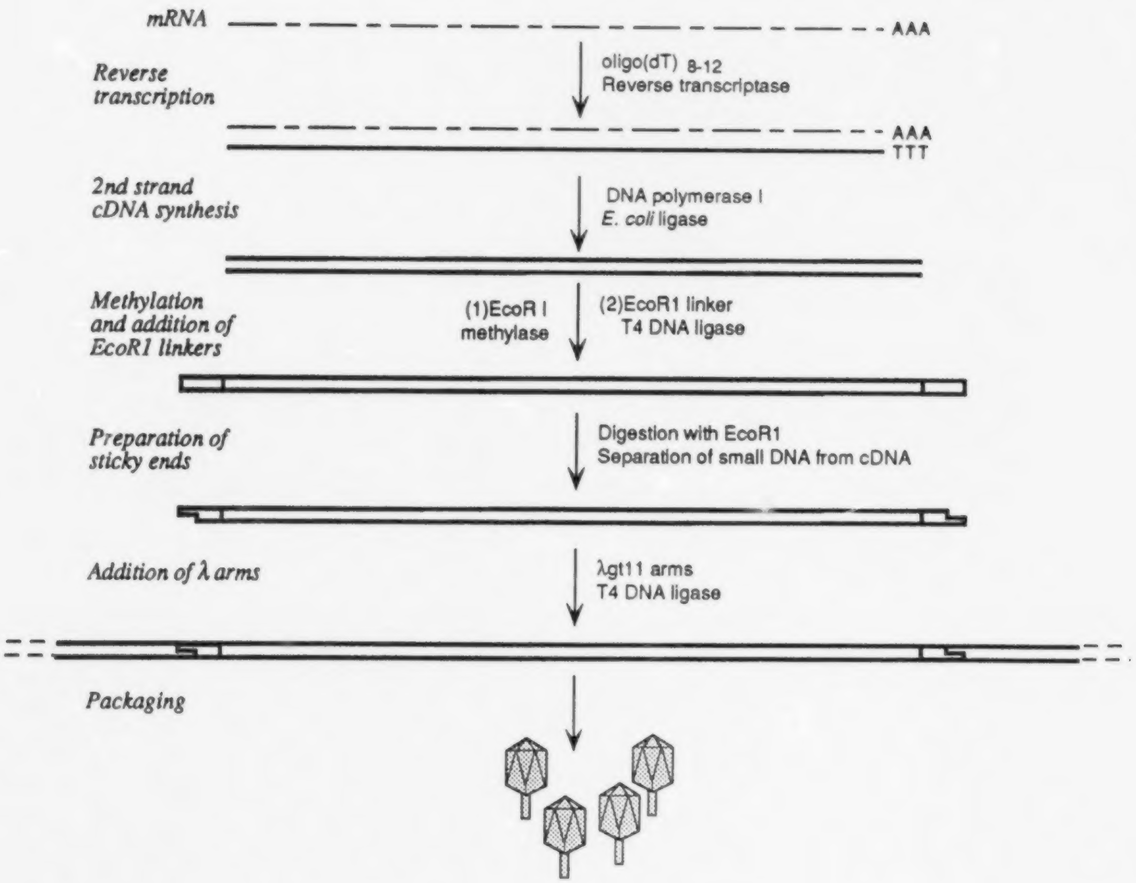
(a) Isolation of total RNA and mRNA

RNA is a chemically stable material under non-alkaline conditions, but it is extremely sensitive to ribonucleases. To ensure the success of the isolation of RNA, efforts were made to eliminate contamination with ribonucleases; these included cleaning all the glassware with water containing diethyl pyrocarbonate (DEPC) before autoclaving and baking; all the plasticware was from new packages and was autoclaved before use; whenever necessary and possible, the solutions and buffers were treated with DEPC solution and autoclaved; all the material involved in RNA preparation was stored away from those for other experiments; gloves were used through the whole operation.

The tissue used for RNA isolation was a heart removed from a 2 day-old piglet after experimental surgery performed by Dr. Coe in the Department of Pediatrics, University of Alberta, Edmonton. Immediately after being removed from the piglet, the heart was homogenized for 60 seconds in about 120ml of a denaturing solution of 4 M guanidinium thiocyanate, 0.1 M Tris-HCl (pH 7.5) and 1% β-mercaptoethanol. After homogenization, the RNA was considered to be safe from degradation because of the presence of guanidinium thiocyanate and β-mercaptoethanol (MacDonald *et al.*, 1987), and the homogenate was transported to the laboratory without any further precautions.

A 60ml homogenate was centrifuged briefly to remove insoluble debris and, after the solution was mixed with 3ml of 2M potassium acetate (pH5.5) and 4.8ml of 1M acetic acid, 48ml of ethanol was slowly added while vortexing. The mixture was stored at -20°C overnight. The resultant precipitate was harvested by centrifugation at 10,000 g for 10

Figure 4. Production of Pig Heart cDNA Library in λgt11


mRNA
AAA
Reverse transcription
oligo(dT) 8-12
Reverse transcriptase
AAA
TTT
2nd strand cDNA synthesis
DNA polymerase I
E. coli ligase
Methylation and addition of EcoR1 linkers
(1)EcoR I methylase
(2)EcoR1 linker
T4 DNA ligase
Preparation of sticky ends
Digestion with EcoR1
Separation of small DNA from cDNA
Addition of λ arms
λgt11 arms
T4 DNA ligase
Packaging

minutes, resuspended in 15 ml of 7.5 M guanidine-HCl (pH7), and precipitated at -20°C overnight after mixing with 0.75 ml of 2 M potassium acetate (pH5.5) and 7.5 ml ethanol. The pellet, after centrifugation at 10,000 g for 15 minutes, was rinsed with ethanol, resuspended in 1 ml of 20 mM EDTA (pH7) and extracted with 24 ml of a mixture of chloroform and n-butanol (4:1/v:v). After centrifugation, the aqueous phase was transferred into a mixture of equal volumes of chloroform and phenol. The organic phase of chloroform and n-butanol was extracted twice more with 0.5 ml 20 mM EDTA (pH7). The pooled aqueous phase, after extraction by phenol/chloroform, was precipitated by 0.1 volume of 2 M sodium acetate (pH7) and 2.5 volumes of ethanol overnight at -20°C. The RNA thus isolated was collected by centrifugation at 10,000 g for 30 minutes and dissolved in water.

About 2 mg of total RNA was obtained from the single pig heart and the RNA appeared to be intact (Figure 5) upon analysis by electrophoresis in an agarose gel containing formaldehyde (see Materials and Methods).

In order to isolate mRNA from total cellular RNA, the "batch" procedure of Jacobson (1987) was used. Oligo(dT)-cellulose (0.3 g) was suspended in 1 ml of binding buffer, consisting of 0.01 MTris-HCl (pH7.5), 0.5 M NaCl, 1mM EDTA, 0.5% SDS, and was centrifuged in a Microfuge for 5 minutes. The cellulose pellet was washed and centrifuged four more times to equilibrate the oligo(dT)-cellulose. A sample of total RNA (0.8 mg) was dissolved in 0.3 ml H_2O, heated at 65°C and then cooled in ice. The RNA was mixed with 0.3 ml of 2x binding buffer of 40mM Tris-HCl (pH7.5), 10mM $MgCl_2$, 20mM ammonium sulfate, 200mM KCl and 100µg/ml BSA and added to the oligo(dT)-cellulose. The poly(A)$^+$ RNA was allowed to bind to the media for 15 minutes with gentle shaking. The bound poly(A)$^+$ RNA was spun down with the cellulose and washed 3 times with binding buffer by suspension and centrifugation. The poly(A)$^+$ RNA was eluted from the oligo(dT)-cellulose 3 times with 0.3 ml of elution buffer consisting of 0.01 M Tris-HCl

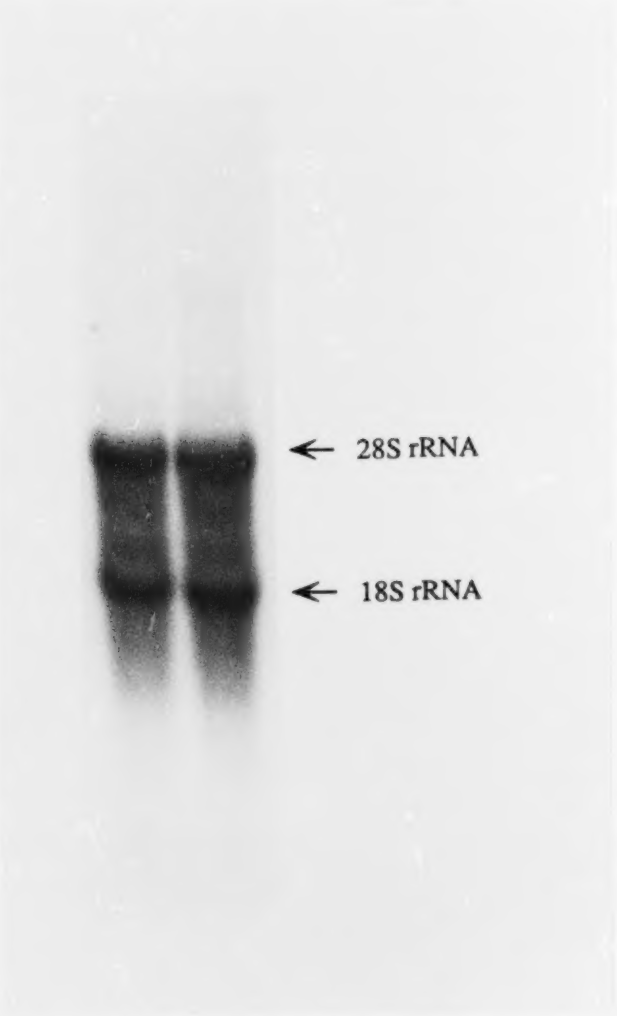


Figure 5. Intact total RNA from pig heart 4 µg of RNA (in two lanes) was analysed on a formaldehyde gel as described by Berger and Kimmel (1967,pp82) without modification. After the electrophoresis the gel was stained in 0.2 µg/ml ethidium bromide and then photographed.

(pH7.5), 1mM EDTA. The eluates were pooled, adjusted to 0.5 M NaCl and then rebound and re-eluted as above. The poly(A)$^+$ RNA in the final eluates was precipitated with ethanol. The yield of this preparation was 2.6% of the total RNA added.

(b) The synthesis of cDNA

The original protocol for the synthesis of cDNA (Gubler and Hoffman, 1983) divided the synthesis of two cDNA strands into two steps with intervening phenol/chloroform extractions and precipitation. I used a modified protocol of cDNA synthesis involving a one-tube reaction scheme which could prevent the lost of cDNA during precipitation and require much less effort (Gubler, 1988).

AMV reverse transcriptase (200 units) was diluted into 10 volumes of 10% glycerol, 10mM potassium phosphate (pH7.4),0.2% Triton X-100, 2mM DTT and incubated in ice for 30 minutes with occasional stirring.

The first cDNA strand was synthesized in 100µl of 50mM Tris-HCl (pH8.3), 10 mM $MgCl_2$, 10mM DTT, 100mM NaCl, 4mM sodium pyrophosphate, 0.625mM each of dATP, dGTP, dCTP, dTTP, 100 µg/ml oligo$(dT)_{12-18}$, and 12 µg poly(A)$^+$ RNA with 200 units of AMV reverse transcriptase at 43°C for 60 minutes. In order to monitor the reaction, 2.5µl of the above mixture was removed and added to an Eppendorf tube containing 1 µCi of α-^{32}P-dCTP and the reaction was carried out as above. To terminate the reaction in the monitoring tube, 2 µl of 0.1 M NaOH, 10mM EDTA was added and the mixture was then boiled for 5 minutes and analyzed by electrophoresis in an alkaline agarose gel (Material and Methods) (Figure 6).

The first strand cDNA synthesis mix was used directly for the second strand cDNA synthesis by the addition of 312.5 µl of 40mM Tris-HCl (pH7.5), 10mM $MgCl_2$, 20mM $(NH_4)_2SO_4$, 200mM KCl, 100 µg/ml BSA, 6.5µl of 15mM β-NAD$^+$, 12 units of

RNase H, 290 units of *E. coli* DNA polymerase I, 6.5 units of *E. coli* ligase and adjusted to a final volume of 625 μl with water. 5 μl of this mixture was removed and added to 1 μCi of α-^{32}P-dCTP in another tube for monitoring. Both tubes were first incubated at 12°C for 60 minutes and then at room temperature for another 60 minutes before being heated at 70°C for 10 minutes. The labeled sample was analyzed on an alkaline agarose gel (Figure 6). To the non-labeled mixture, 25 units of T4 DNA polymerase was added and a reaction to generate blunt-ends on the cDNA was carried out for 15 minutes at 37°C. The mixture containing the cDNA was extracted twice with phenol after it was adjusted to contain 20mM of EDTA and 1% of SDS and the cDNA was then precipitated with 0.1 volume of 3 M sodium acetate and 2 volumes of ethanol at -70°C. The cDNA was reprecipitated and the pellet formed by centrifugation was rinsed with 70% ethanol, dried and dissolved in 11 μl of water. 1 μl of the cDNA was used for measurement of concentration, which indicated that there was about 5 μg of cDNA in the remaining 10μl.

The cDNA synthesized in the presence of α-^{32}P-dCTP was analyzed on an alkaline agarose gel, as described above, under conditions where RNA would be degraded. Figure 6 is the autoradiograph of the gel after being dried. Judging from the ^{32}P end-labeled and denatured λ DNA fragments digested by Hind III, the sizes of the both strands of the cDNA ranged from 500 bp to longer than 4,000 bp.

(c) Generation of EcoR I sites on the ends of the cDNA

In order to be cloned into λgt11, EcoR I sites must be created at both ends of the cDNA. This was done by ligation of synthetic EcoR I linkers to the ends of the cDNA followed by digestion with EcoR I. To prevent the cDNA from cleavage by EcoR I, internal EcoR I sites were protected by methylation before the addition of EcoR I linkers.

5 μl of the cDNA was methylated in 100 mM Tris-HCl (pH 8), 1 mM EDTA, 0.16 mM S-adenosylmethionine, 0.2 mg/ml BSA, and 20 units of EcoR I methylase in a

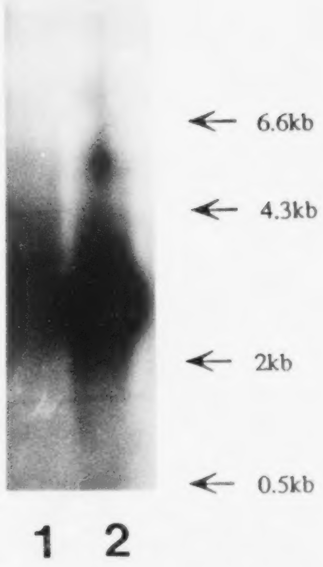


Figure 6. Synthesis of cDNA The cDNA was analysed in an alkaline agarose gel as descibed by Berger and Kimmel (1987, pp79) without modification. After the electrophoresis, the gel was dried and autoradiographed. Lane 1: the products of the first strand cDNA synthesis; Lane 2: the products of the seconf strand cDNA synthesis. The arrows indicate the sizes of the standard DNA fragments. The radioactive small molecules were still in the gel, so the bottom of the autoradiograph was over exposed.

final volume of 20 µl and incubated at 37°C for 90 minutes. The methylated cDNA was extracted first by phenol/chloroform and then by ether saturated with water before being precipitated by ethanol. As a control, the λ DNA fragments produced by Hind III digestion were also methylated in the same manner and were found to be protected fully by the action of methylation and thus made resistant to digestion by EcoR I.

The methylated cDNA was ligated with 1.6 µg EcoR I linkers in a volume of 15 µl of 66 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 10 mM DTT, 0.45 mM ATP, 1 mM spermidine, 0.2 mg/ml BSA and 1.5 unit T4 DNA ligase in 14°C for 40 hours. The amount of EcoR I linkers used was at least 40-fold in molar excess over the cDNA in order to minimize the frequency of ligation between cDNA molecules. This ligation mix was heated at 70°C for 10 minutes to inactivate the ligase and then was drop-dialyzed (Marusyk, 1980) against 10 mM Tris-HCl (pH 7) and 1 mM EDTA before digestion with 360 units of EcoR I at 37°C overnight in a volume of 50 µl of 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM NaCl and 1 mM dithioerythritol. The sample thus treated was extracted with phenol and chloroform and then with ether before sizing.

(d) Separation of cDNA from the free linkers and the completion of the library

Before ligation of the cDNA to the vectors, the free linkers in the sample must be separated from the cDNA, otherwise their dominant concentration would make the vectors unavailable to the cDNA. Traditionally, this has been done by gel filtration chromatography using a column that could be as large as 32 cm in length and 0.2 cm in inner diameter packed with either Bio-Gel A-50m (Bio-Rad), Ultrogel AcA 34 (LKB) or Sepharose CL-4B (Pharmacia) (Huynh *et al.*, 1985). We found that this procedure could be replaced by chromatography using an FPLC instrument equipped with a Superose™ 12 column, based upon a brief report of using HPLC for the same purpose (White *et al.*, 1988). The FPLC chromatography was performed with the assistance of Mrs. Shirley Shechosky, as follows.

A Superose™ 12 column was equilibrated with 10mM Tris-HCl (pH8), 1mM EDTA and standardized with DNA fragments of pBR322 digested with HinfI and λ DNA digested with Hind III and EcoRI. The elution rate was set at 0.5 ml/min and the fraction size was set at 0.5ml. A typical experiment could be completed in less than 45 minutes. The progress of the chromatography was monitored by measurement of absorbance at 254nm, but it was found that the instrument lacked the sensitivity required to follow the amount of DNA present and thus it was not used as the means to trace the DNA. Instead, each fraction of the eluate was precipitated and electrophoresed in PAGE buffered by TBE (Material and Methods) to identify the fractions containing the DNA standards (Figure 7). To our satisfaction, the elution profiles of FPLC equipped with Superose™ 12 column were very reproducible and, after several trial experiments, the cDNA was applied to the column and the fractions expected to contain the desired cDNA were collected without further complication of the traditional isotope tracing. The fractions expected to contain the desired cDNA were pooled and found to comprise a total of 129 ng of DNA. The cDNA was co-precipitated with 7 μg dephosphorylated λgt11 arms by ethanol. After pelleting and resuspension in TE buffer, the DNA was ligated and packaged *in vitro* using Packgene™ from Promega and a DNA Packaging Kit from Boehringer Mannheim according to the manufacturers' instructions.

The library was either screened immediately or amplified (Huynh *et al.*, 1985) for later use.

(4) RESULTS AND DISCUSSION

There are several criteria that can be used to judge the qualities of a cDNA library. A good cDNA library should reflect the population of mRNA. There are, normally, 1×10^4 to

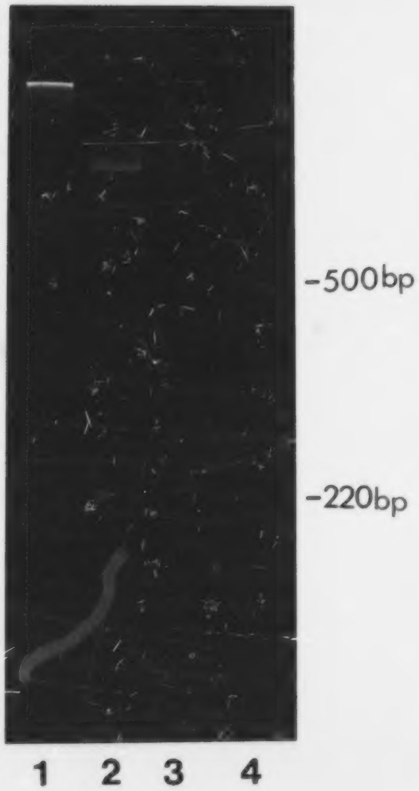


Figure 7. Fractions from FPLC chromatography The DNA in fractions from FPLC chromatography were precipitated by ethanol. The precipitates, after being redissolved in water, were analyzed on PAGE buffered by TBE. Lanes 1 to 4 were loaded with samples corresponding to that eluted from 6.5 to 8.5 ml. The DNA standards used were mixtures of pBR322 DNA digested by Hinf I and λ DNA digested by EcoR I and Hind III.

3×10^4 mRNA species in a vertebrate cell, so there should be about 10^6 independent recombinant clones in a representative cDNA library to ensure that rare mRNA is likely to be included (Kimmel and Berger, 1987; pp314). Moreover, the clones in the library should closely represent the contents of mRNA — inserts should be large enough to include entire messages.

Several experiments were carried out to test the newly constructed library against the above criteria.

Firstly, aliquots of the library were plated out on LB/ampicillin plates containing IPTG and x-gal. The library was thus determined to be able to produce 1×10^6 independent plaques. Among these plaques, more than 98% of them did not produce blue color, indicating that their coding regions for the α-peptide of β-galactosidase, which was responsible for the color production, had been interrupted by a the presence of a foreign gene and thus were recombinants.

Because neither λgt10 nor λgt11 need 'stuffer' DNA to yield packagable DNA, a library constructed by these vectors can comprise 70-99% non-recombinants (Huynh et al, 1985, pp70) without special treatments. Most commercially available libraries of λgt11 contain 70-80% non-recombinants and require screening of a large number of plaques to obtain a particular clone. This problem is eliminated in λgt10 by the presence of a cI gene ($cI^{+}imm^{434}$). The cI gene in λgt10 is a λ repressor gene with a unique EcoR I site. Insertion of foreign DNA at this restriction site interrupts the cI gene and causes the phenotype of the phage to change from cI^{+} in wild type to cI^{-} in recombinant. Therefore, when an amplification is carried out by infection of an *E. coli* strain carrying the high frequency lysogeny mutation, hflA150, the $cI^{+}imm^{434}$ λgt10 phages of non-recombinants are efficiently repressed and do not lyse the cells. The insert-containing $cI^{-}imm^{434}$ λgt10 phages, in contrast, can lyse the hflA150 strains normally and would thus be the major component in the amplified library (Huynh *et al.*, 1985). λgt11 does not contain a genetic selection marker, even though it has the genetic indicator β-galactosidase, and the only way to eliminate the large numbers of non-recombinants is to dephosphorylate the λgt11 arms before ligation to prevent ligation between the λgt11 arms. But there is a drawback to this treatment: it was found that the phosphatase-treatment of the λgt11 arms could result in a substantial 70-fold decrease in the yield of total recombinants (Lapeyre and Amalric, 1985). Besides the problem of obtaining both high percentage of recombinants and high titer of the λgt11 libraries at the same time, it was stated that one normally requires 10 to 40 μg mRNA as starting material to construct a library containing 10^6 clones (Kimmel and Berger, 1987, pp315). These factors indicated that the procedure used in the generation of this library is efficient since 10^6 clones were generated from 6 μg of mRNA.

Secondly, 12 plaques were randomly selected for λ DNA purification. The inserts in these λ DNA's were excised and separated on a agarose gel (Figure 8). As indicated, all the clones contain inserts. Judging from the DNA standards run along side, the inserts ranged from about 800 bp to 4,000 bp with an average size of 1,600bp.

As mentioned above, the library comprised mostly phage that produce white plaques when grown on x-gal plates and should therefore contain inserts. However, white plaques do not mean useful clones if the inserts are very small. The data derived from these randomly selected clones suggested that this library could be a good source of full length clones for pig heart proteins; the average size of inserts of typical cDNA libraries is about 800 bp or smaller (Huynh *et al.*, 1985, p71).

Lastly, the most important criterion of a good cDNA library is the ultimate success in cloning the desired cDNA. As will be described later in this thesis, several cDNA clones encoding CoA transferase and the α subunit of pig heart SCS were satisfactorily isolated from this library using either a cDNA clone from another source or synthetic

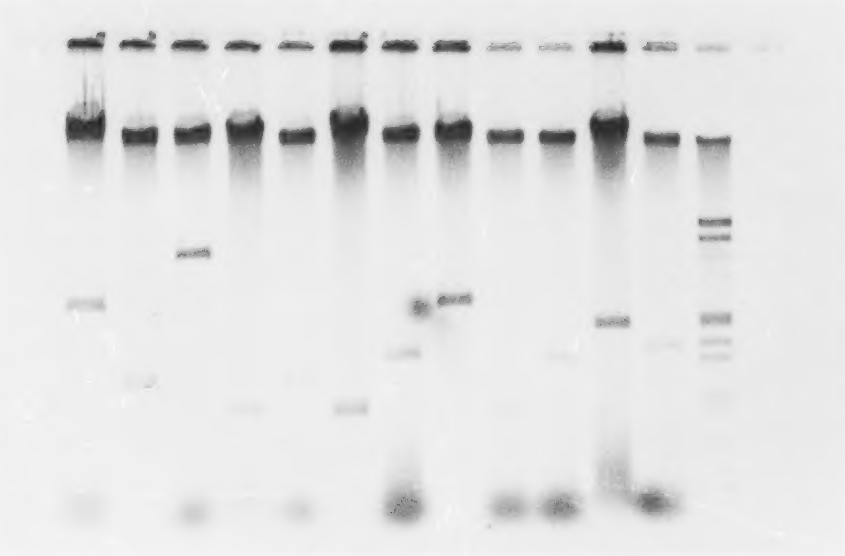


Figure 8. Inserts of the randomly picked clones The λgt11 DNA was digested with EcoR I and then separated on a 0.8% agarose gel buffered by TAE (Berger and Kimmel, 1987, pp78). Lane 1-12: DNA from randomly selected clones, digested with EcoR I. Lane St: DNA standards of λ DNA digested by EcoR I and Hind III

oligonucleotides as probes. Moreover, the clone of the full length cDNA of the β subunit of SCS was also successfully isolated from this library by Darrin Bailey in the laboratory using antibodies as probes.

CHAPTER IV

cDNA CLONING OF THE PRECURSORS OF THE α SUBUNITS OF PIG HEART SUCCINYL-COA SYNTHETASE

(1) INTRODUCTION

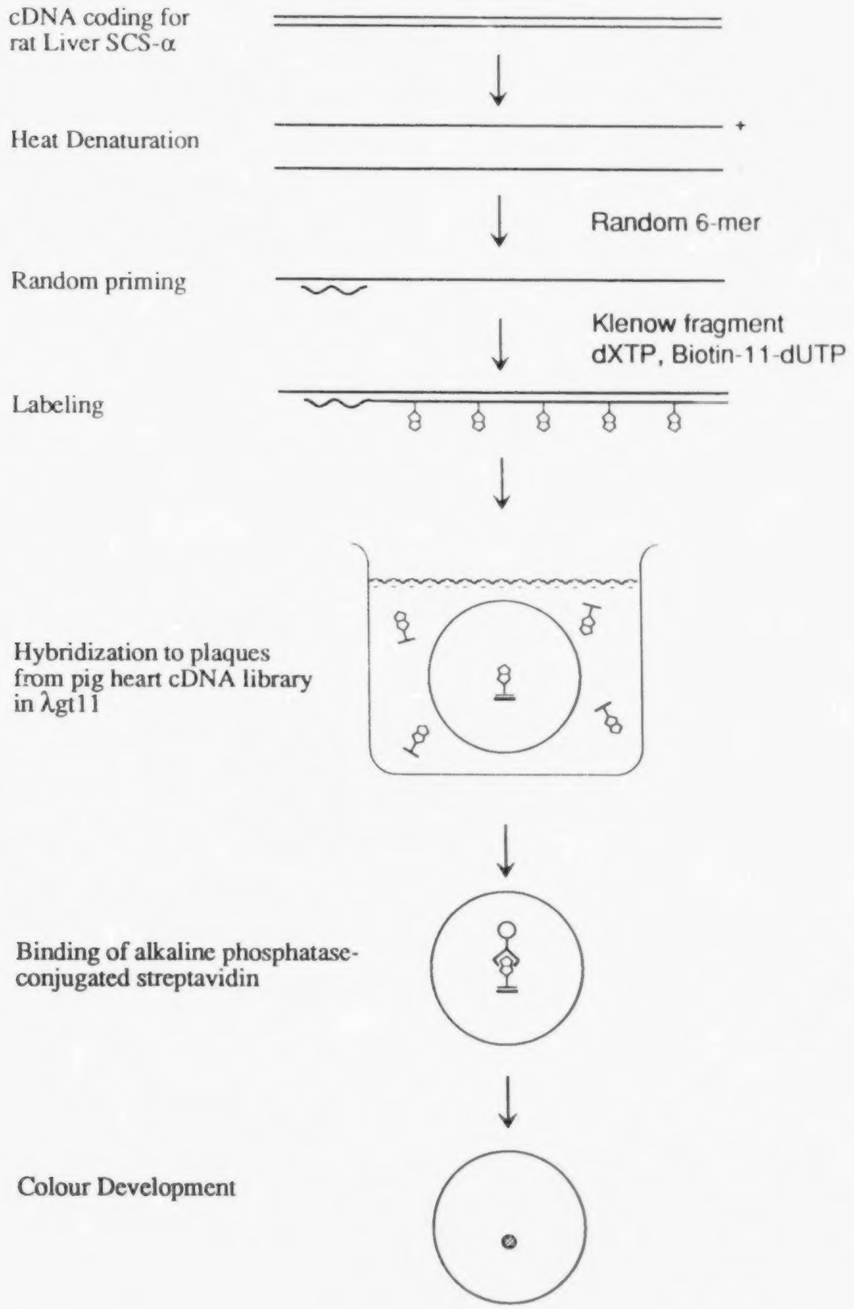
A striking feature of the α subunits of SCS of rat liver and *E. coli* (the α subunits of rat liver and *E.coli* were the only ones that had been studied by cDNA cloning technology before this projects started) was the homology between the two species. It had been shown that there was ~70% identity between the amino acid sequences translated from the two genes with virtually no gapping (Henning *et al.*, 1988). It could be speculated that the homology between two eukaryotic cDNA's coding for the α subunits of SCS would be even higher, as might be revealed from the sequences of the rat liver and pig heart proteins. If this was true, the available cDNA coding for the α subunit of rat liver SCS could be turned into a very promising probe to clone the cDNA encoding the α subunit of pig heart SCS. However, in the absence of information on the homology and without knowledge of tightness of binding of the rat liver cDNA to that of pig heart, a strategy was planned to clone the pig heart cDNA using the probes derived from the rat liver cDNA in a series of hybridization experiments with gradually increasing stringencies.

The recent development of systems for labeling DNA with biotin-avidin (streptavidin) coupled with a colour-generating enzyme reaction instead of radioactive labeling of DNA probes (Langer *et al.*, 1981; Leary *et al.*, 1983) offered an alternative for the generation of probes. There are obvious advantages to using the non-radioactive probe: it avoids the hazards of radioisotopes; the probes are stable and, once made, can be used again and again or stored for a long period of time; the reagents are relatively inexpensive and have extremely long lifetimes compared to that of isotopes; and the method offered a sensitivity that is almost as good as that of radioactive probes (in our hands).

The basis of non-isotope labelling for the generation of DNA probes is that an analogue of dTTP (derived from dUTP by attaching biotin at the 5 -carbon of uracil through a spacer arm) could be incorporated into DNA by *E. coli* DNA polymerase I in place of dTTP, and the biotin thus incorporated could be very tightly bound by avidin or streptavidin. Once the biotin-labelled DNA is hybridized to a target DNA, it can be detected by colour reactions of alkaline phosphatase coupled to avidin or streptavidin (Figure 9).

There are several commercially available non-isotopic labeling systems. However, on close examination it was found that there were a number of problems with each of these systems. Let us consider the DIG DNA labeling system from Boehringer-Mannheim as a example. The principle of this kit is that DIG (digoxigenin, a compound that had been used as hapten to raise monoclonal antibodies) conjugated dUTP could be incorporated into DNA through labeling reactions and, once the DIG-dUTP containing probe bound to the target DNA, it could be detected by the colour reaction of phosphatase conjugated to the monoclonal antibody against DIG (see brochure from Boehringer-Mannheim). The major drawback of this kit was that it had low sensitivity. The manufacturer offered suggestions to tackle that problem, but these were only at the expense of producing possibly undesirable side-effects. For example, it was suggested to label the DNA for up to 20 hours to increase the amount of incorporated probe. It was discovered that the Klenow fragment of DNA polymerase I could produce non-template related DNA, if one lets the polymerization reaction proceed for more than 6 hours (Radding *et al.*, 1962). Therefore, DNA probes generated from extended labelling reaction of random priming could produce non-specificity in the probing. This was not recommended. The system from this company also required very high probe concentration (200 ng/ml) and long times (as long as 24

Figure 9. Screening Pig Heart cDNA Library with Biotinylated Probes


cDNA coding for rat Liver SCS-α
Heat Denaturation
+
Random 6-mer
Random priming
Klenow fragment
dXTP, Biotin-11-dUTP
Labeling
Hybridization to plaques from pig heart cDNA library in λgt11
Binding of alkaline phosphatase-conjugated streptavidin
Colour Development

hours) for color development. Another kit called BluGENE™ offered by Bethesda Research Laboratories had similar problems.

Unsatisfied with the commercial kits, we decided to combine various protocols (Haas and Fleming, 1986, 1988; Kincaid and Nightingale, 1988) with our knowledge to design our own experiments and choose reagents we believed to be the best or the least expensive without compromising the quality. Although a direct comparison was not done, I believe that the procedure described below could produce better results with simplicity and better economy to those provided by the commercial kits and protocols of those companies.

(2) EXPERIMENTAL PROCEDURES

The DNA labeling method was derived from the now classic random priming methods (Feinberrg and Vogelstein, 1983, 1984) but used the biotin-11-dUTP or biotin-18-dUTP in the place of α-[^{32}P]-ATP of the model protocol. The random priming method has been proven to be superior to the nick translation methods in labelling the DNA to high specific radioactivity.

Linear DNA (20 to 100 ng, normally isolated from agarose gels (see Material and Methods) was heated at 95° to 100°C for 10 minutes, then quickly cooled in ice. For a final volume of 50 µl, the denatured DNA was mixed with 20 µM dATP, dGTP, dCTP, dTTP, 50 mM Tris-HCl (pH 8), 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, 0.2 mg/ml BSA, 200 mM HEPES (pH 6.6), 20 µM biotin-11-dUTP or biotin-18-dUTP, 5 OD/ml hexanucleotides and 2 units of the Klenow fragment of *E. coli* DNA polymerase I. The reaction was allowed to proceed at room temperature for 6 hours and the sample was heated at 70°C for 10 minutes to stop the reaction. The DNA thus labeled could be used directly or stored at -20°C.

The *E. coli* host Y1088 was transfected with the pig heart cDNA library in λgt11 and was spread onto LB/ampicillin plates (85 mM diameter) with top agarose (7% agarose prepared in LB medium) at a density of 1 to 1.5x10^4 pfu. Phage replication was allowed to proceed at 37°C until the plaques just started to contact one another. The plates were chilled at 4°C for at least 1 hour. A piece of nitrocellulose paper was placed onto the surface of the lysed bacterial lawn for 30 seconds (1 minute for a duplicate) and then peeled off. The nitrocellulose, with plaque-absorbing side up, was laid onto a sheet of Whatman 3MM paper soaked with 0.5 M NaOH and 1.5 M NaCl for 5 minutes, then transferred to another sheet of Whatman 3MM paper soaked in 0.5 M Tris-HCl (pH 8) and 1.5 M NaCl for 5 minutes before being immersed in the solution of 0.5 M Tris-HCl (pH 8) and 1.5 M NaCl for 5 minutes and 2xSSC (1xSSC is 0.15 M NaCl, 0.015 M trisodium citrate (pH 7)) for 5 minutes. The filters were air-dried and baked at 80°C for 2 hours *in vacuo*.

After being wetted with TBS (50 mM Tris-HCl (pH 7.5) and 150 mM NaCl), the filters were treated with proteinase K at a concentration of 100 μg/ml in TBS supplemented with 10 mM $MgCl_2$ for 60 minutes at 37°C. The protease-treated filters were, after being rinsed in TBS, pre-hybridized twice, for half an hour each time, in a pre-hybridization solution of 5xSSC supplemented with 25 mM sodium phosphate (pH 6.5), 5x Denhardt's solution (1 x Denhardt's solution is 0.02% BSA, 0.02% Ficoll, 0.02% polyvinyl-pyrrolidone), 0.5 mg/ml denatured sheared herring sperm DNA. The filters then were hybridized with the freshly denatured biotinylated probes at a concentration of as low as 1 ng/ml (could be up to 20 ng/ml) in 5% dextran sulfate and pre-hybridization solution overnight. In a typical experiment, 30 filter circles were shaken in 60 ml of hybridization solution in a beaker covered with aluminium foil. After the experiment, the hybridization solution was stored at -20°C. Probes generated from the cDNA isolated from a mini-prep of double-stranded M13 DNA were thus used to process more than 100 filters in several rounds of experiments without losing signal intensity. The filters were washed twice with

2xSSC and 0.1% SDS at room temperature, and twice with 0.1xSSC and 0.1% SDS at the hybridization temperature before being rinsed in 0.1xSSC and air-dried.

The hybridization temperature was set at 40°C for the starting low stringency experiments and then gradually increased to 65°C for the final round of hybridization. The hybridization solution once made was used throughout the whole process with heating and cooling before each experiment.

For colour development, the filters were rehydrated in 0.1 M Tris-HCl (pH7.5) and 0.15 M NaCl for 1 minute, then shaken in 3% BSA, 0.1 M Tris-HCl (pH7.5), 0.15 M NaCl (re-usable solution, stored at -20°C) for 1 hour, before being either air-dried or dried under vacuum at 80°C for 10 to 20 minutes. The filters were rehydrated again, this time with 3% BSA in 0.1 M Tris-HCl (pH 7.5) and 0.15 M NaCl, and were then incubated with streptavidin conjugated with alkaline phosphatase (BRL) at a 2000-fold dilution in 0.1 M Tris-HCl (pH7.5) and 0.15 M NaCl for 1 hour at room temperature. The filters were then washed twice in large volumes of 0.1 M Tris-HCl (pH7.5) and 0.15 M NaCl at room temperature, 15 minutes each time. The filters were next washed in a solution of 0.1 M Tris-HCl (pH9.5), 0.1 M NaCl and 50 mM $MgCl_2$ for 10 minutes before the colour was developed (in the dark) in the same solution plus 3.3 mg/ml nitroblue tetrazolium (NBT) and 1.7 mg/ml 5-bromo-4-chloro-3-indolylphosphate (BCIP) for 30 minutes. The colour reaction was terminated by rinsing the filters in 20 mM Tris-HCl (pH7.5) and 0.5 mM EDTA, then the filters were air-dried. The plaques that had the complementary sequences with the probes could be identified by alignment with the dots of the bluish purple color for the positive signals in the filters with the plaques on the plates.

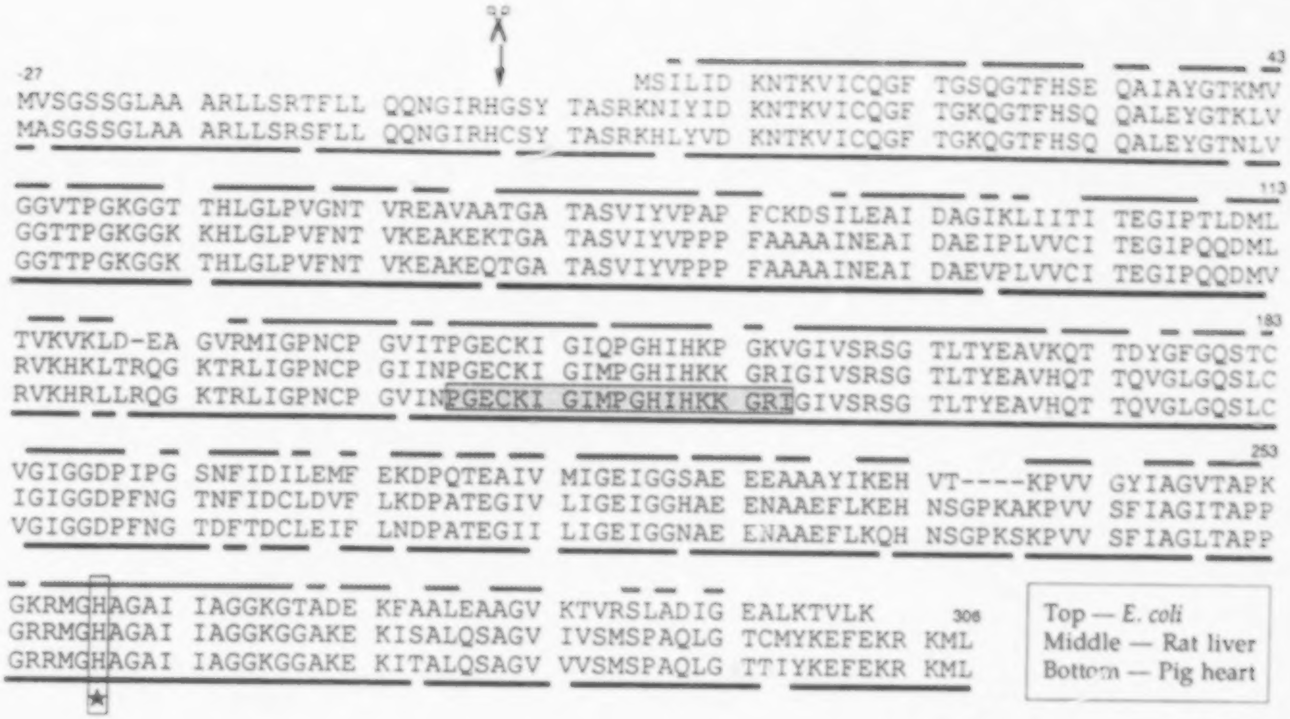
Positive plaques were selected, the λ DNA was isolated (Maniatis *et al.*, 1982, pp373; Grossberger, 1987; Helms *et al.*, 1987) and the inserts were digested from the vector by EcoR I and separated on agarose gels along with DNA standards. The DNA fragments of the inserts were subcloned into M13mp18 and sequenced by the dideoxynucleotide termination method (Sanger *et al.*, 1977) using Sequenase™ and Sequenase™ version II, which were T7 DNA polymerases modified to eliminate the 5' to 3' exonuclease activity by either chemical methods or genetic engineering (Tabor and Richardson, 1987, 1989).

(3) RESULTS AND DISCUSSION

Both amplified and unamplified versions of the pig heart cDNA library were screened. More than 50 clones positive to the probes of the cDNA of the α subunit of rat liver SCS were identified from the amplified library. Ten clones were also identified as positive to the probes from about 1×10^4 plaques of the unamplified library, and these were all later confirmed to be the desired clones by means of DNA sequencing or polymerase chain reaction. Moreover, all of these ten clones contained the full coding sequence, from the sequence upstream of the signal peptide to the poly(A)$^+$ tail, additional evidence in support of the exceptional quality of the pig heart library that was prepared.

Screening for the clone coding for the α subunit of pig heart SCS originally was intended as a side project of the cloning of CoA transferase, and the generation of the cDNA library was also planned to pave the path for the cloning of the β subunit of SCS. The task was expected to be relatively straightforward and the cDNA's cloned by our approach were expected to have basically the same coding sequence as that of the α subunit of SCS from rat liver. However, as sequencing data became available, the unexpected happened: it was discovered that there were not only clones of expected sequences but positive clones having unexpected content as well.

The nucleotide sequence of one of the clones (PS57) was found to be very similar with that coding for the α subunit of rat liver SCS in the region of the longest reading frame (Figure 10) with a derived amino acid sequence that was virtually identical to that of the α



```
-27                                                                                                    43
                                           MSILID KNTKVICQGF TGSQGTFHSE QAIAYGTKMV
MVSGSSGLAA ARLLSRTFLL QQNGIRHGSY TASRKNIYID KNTKVICQGF TGKQGTFHSQ QALEYGTKLV
MASGSSGLAA ARLLSRSFLL QQNGIRHCSY TASRKHLYVD KNTKVICQGF TGKQGTFHSQ QALEYGTNLV

                                                                                                      113
GGVTPGKGGT THLGLPVGNT VREAVAATGA TASVIYVPAP FCKDSILEAI DAGIKLIITI TEGIPTLDML
GGTTPGKGGK KHLGLPVFNT VKEAKEKTGA TASVIYVPPP FAAAAINEAI DAEIPLVVCI TEGIPQQDML
GGTTPGKGGK THLGLPVFNT VKEAKEQTGA TASVIYVPPP FAAAAINEAI DAEVPLVVCI TEGIPQQDMV

                                                                                                      183
TVKVKLD-EA GVRMIGPNCP GVITPGECKI GIQPGHIHKP GKVGIVSRSG TLTYEAVKQT TDYGFGQSTC
RVKHKLTRQG KTRLIGPNCP GIINPGECKI GIMPGHIHKK GRIGIVSRSG TLTYEAVHQT TQVGLGQSLC
RVKHRLLRQG KTRLIGPNCP GVINPGECKI GIMPGHIHKK GRIGIVSRSG TLTYEAVHQT TQVGLGQSLC

                                                                                                      253
VGIGGDPIPG SNFIDILEMF EKDPQTEAIV MIGEIGGSAE EEAAAYIKEH VT----KPVV GYIAGVTAPK
IGIGGDPFNG TNFIDCLDVF LKDPATEGIV LIGEIGGHAE ENAAEFLKEH NSGPKAKPVV SFIAGITAPP
VGIGGDPFNG TDFTDCLEIF LNDPATEGII LIGEIGGNAE ENAAEFLKQH NSGPKSKPVV SFIAGLTAPP

GKRMGHAGAI IAGGKGTADE KFAALEAAGV KTVRSLADIG EALKTVLK      306
GRRMGHAGAI IAGGKGGAKE KISALQSAGV IVSMSPAQLG TCMYKEFEKR KML
GRRMGHAGAI IAGGKGGAKE KITALQSAGV VVSMSPAQLG TTIYKEFEKR KML
```

Top — *E. coli*
Middle — Rat liver
Bottom — Pig heart

Alternatively spliced cassette: RFLRFAVYISRVFTCTQQEEYRKIPLLFGTRSIHM

Figure 10. Comparison of the predicted amino acid sequences of the α subunits of SCS The numbering is based on the rat liver sequence. The arrow with the scissors indicates the point of proteolytic cleavage to remove the signal sequence for the eukaryotic proteins. Using the rat liver sequence as the standard, homology is indicated as follows: the upper shaded line indicates identity between *E.coli* and rat liver subunits; the lower solid line indicates identity between pig heart and rat liver subunits. The shaded blocks indicate the sequence of alternative spliced cassette. The box with the star indicates the position of his259, the active site histidine residue.

subunit of rat liver SCS and about 70% identical to that of the α subunit of *E. coli* SCS (Figure 11). It seemed that there was no doubt that this clone encoded the α subunit of pig heart SCS and that it was the cDNA that we intended to clone.

The nucleotide sequence of another clone (PS108) had the unexpected content. It had the almost identical nucleotide sequence to that of PS57, including the non-coding sequences at both 5' and 3' ends, except that a stretch of 108 base pairs in the middle of the coding sequence replaced a stretch of 57 base pairs of PS57 (Figure 12). Furthermore, this 108 nucleotide sequence did not interrupt the reading frame, but instead it encoded a different sequence of 36 amino acid residues, compared to that of PS57 which coded for a sequence of 19 amino acid residues. This new sequence of PS108 had no counterpart in the clones that had then been isolated for the α subunits of *E. coli*, rat liver, nor in the clones subsequently isolated from yeast (Reed, *et al.*, 1989) and *Thermus aquaticus* (Nicholls, *et al.*, 1988).

This PS108 clone may have represented a minor species of mRNA since it was the only clone containing the unique 108 nucleotide sequence (108-cassette) among the 12 clones examined. All the other clones probably contained 57-cassettes based on DNA sequencing, and polymerase chain reaction experiments using the primers derived either from the 57-cassette or 108-cassette paired with a primer derived from the downstream sequence of the cassettes (Figure 13).

The unexpected discovery of the new sequence of PS108 raised the question of its authenticity. AMV reverse transcriptase does not possess the editing capability associated with a 3' to 5' exonuclease activity and it is not uncommon for it to misincorporate nucleotides during cDNA synthesis. It has been stated that reverse transcriptase could make a mistake statistically around every 500 bp in high concentrations of deoxynucleoside triphosphates and Mn^{++} (Sambrook, *et al.*, 1989, pp5.53).

```
                                   GAATTCCCCGCAGGCCTCGTCGCCGCCCCGGCCGCCGCCACC

ATG GCC TCT GGC AGC AGC GGC CTT GCC GCC GCC CGC CTC CTC TCG CGC TCC TTC CTC TTG
met ala ser gly ser ser gly leu ala ala ala arg leu leu ser arg ser phe leu leu

CAA CAG AAT GGA ATT CGG CAT TGT TCC TAT ACA GCT TCC CGG AAG CAT CTC TAT GTT GAT
gln gln asn gly ile arg his cys ser tyr thr ala ser arg lys his leu tyr val asp

AAA AAT ACG AAG GTT ATT TGC CAG GGT TTC ACT GGT AAG CAG GGC ACC TTT CAT AGC CAG
lys asn thr lys val ile cys gln gly phe thr gly lys gln gly thr phe his ser gln

CAG GCG CTG GAA TAT GGC ACC AAT CTA GTT GGA GGA ACC ACT CCA GGG AAA GGA GGC AAG
gln ala leu glu tyr gly thr asn leu val gly gly thr thr pro gly lys gly gly lys

ACG CAT CTG GGC CTA CCA GTC TTT AAT ACT GTG AAG GAG GCC AAA GAA CAA ACC GGA GCG
thr his leu gly leu pro val phe asn thr val lys glu ala lys glu gln thr gly ala

ACG GCA TCT GTC ATT TAT GTC CCT CCC CCT TTT GCT GCT GCT GCC ATT AAT GAA GCC ATC
thr ala ser val ile tyr val pro pro pro phe ala ala ala ala ile asn glu ala ile

GAT GCA GAA GTG CCC TTG GTC GTG TGC ATC ACT GAA GGC ATC CCG CAG CAG GAC ATG GTG
asp ala glu val pro leu val val cys ile thr glu gly ile pro gln gln asp met val

CGG GTA AAG CAC AGG CTG CTG CGC CAG GGA AAG ACG AGG CTG ATC GGG CCA AAC TGC CCT
arg val lys his arg leu leu arg gln gly lys thr arg leu ile gly pro asn cys pro

GGA GTC ATC AAT|CCT GGA GAA TGC AAA ATT GGC ATC ATG CCT GGC CAT ATT CAC AAG AAA
gly val ile asn|pro gly glu cys lys ile gly ile met pro gly his ile his lys lys

GGA CGA ATT|GGT ATC GTG TCC AGA TCT GGC ACC CTG ACT TAT GAA GCA GTT CAT CAA ACA
gly arg ile|gly ile val ser arg ser gly thr leu thr tyr glu ala val his gln thr

ACA CAA GTT GGC TTG GGG CAG TCT TTG TGT GTT GGC ATT GGA GGT GAT CCT TTC AAT GGA
thr gln val gly leu gly gln ser leu cys val gly ile gly gly asp pro phe asn gly

ACA GAT TTT ACT GAT TGC CTT GAA ATC TTT CTG AAT GAT CCA GCC ACA GAA GGC ATC ATA
thr asp phe thr asp cys leu glu ile phe leu asn asp pro ala thr glu gly ile ile

TTG ATT GGT GAA ATT GGT GGT AAT GCG GAA GAG AAT GCT GCA GAA TTT TTG AAG CAA CAT
leu ile gly glu ile gly gly asn ala glu glu asn ala ala glu phe leu lys gln his

AAT TCA GGT CCC AAG TCC AAG CCG GTG GTG TCC TTC ATT GCT GGC TTA ACT GCC CCT CCT
asn ser gly pro lys ser lys pro val val ser phe ile ala gly leu thr ala pro pro

GGC AGA AGA ATG GGT CAC GCA GGG GCA ATT ATT GCT GGA GGA AAG GGA GGC GCT AAA GAG
gly arg arg met gly his ala gly ala ile ile ala gly gly lys gly gly ala lys glu

AAG ATC ACT GCC CTG CAG AGT GCA GGA GTT GTG GTC AGC ATG TCT CCT GCA CAG CTG GGA
lys ile thr ala leu gln ser ala gly val val val ser met ser pro ala gln leu gly

ACC ACC ATC TAC AAG GAG TTT GAA AAG AGG AAG ATG CTA TGA AAGCAAAACGAATGAGA
thr thr ile tyr lys glu phe glu lys arg lys met leu ***
```

Alternative Block:

```
AGA TTT CTT AGA TTT GCT GTA TAT ATA ACA TCC AGG GTT TTT ACT TGT ACT CAG CAG GAA
arg phe leu arg phe ala val tyr ile thr ser arg val phe thr cys thr gln gln glu

GAA TAC AGG AAG ATA CCT CTG CTT TTT GGA ACC AGA AGT ATC CAT ATG
glu tyr arg lys ile pro leu leu phe gly thr arg ser ile his met
```

Figure 11. Nucleotide and amino acid sequences of PS57 and PS108
The main sequence given is that derived from PS57. The shaded block indicates the 57-base block that is replaced by an alternative sequence of 108 nucleotides in PS108 (sequence shown at bottom).

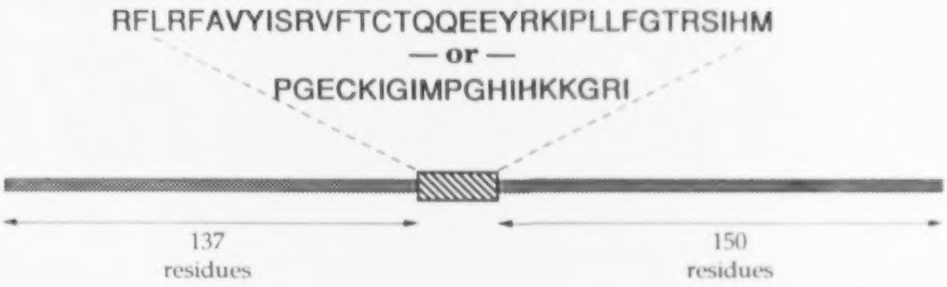

RFLRFAVYISRVFTCTQQEEYRKIPLLFGTRSIHM
— or —
PGECKIGIMPGHIHKKGRI
137
residues
150
residues

Figure 12. Schematic diagram of the two protein isoforms that would be produced by alternative splicing of cassette exons

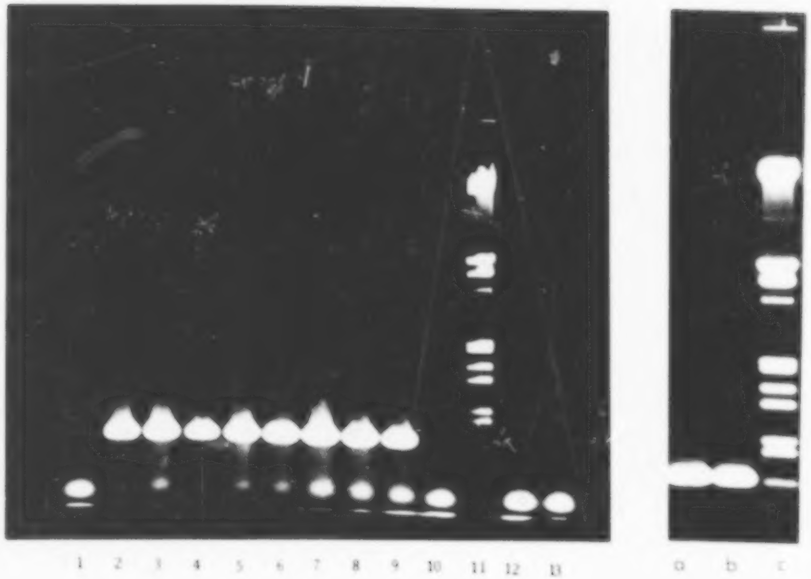


Figure 13. Identification of the clones coding for the α subunits of SCS with 57 cassettes. The PCR was performed according to Saiki *et al.* (1988) with [$MgCl_2$] of 1mM. The products were separated in an agarose gel. Wells 1 to 10 are the PCR products using λPS11 to 20 phage as templates and 57-cassette specific primers. The less intense bands from samples of λPS11 and 20 were due to the limited amount of templates present. Wells 12 and 13 are the PCR products from 108-cassette specific primers using λPS11 and 20 as templates. Wells a and b are the same PCR products of the 108-cassette specific primers using λPS108 as templates. Well 11 and c are λ DNA digested with EcoRI and Hind III.

However, it was felt that this argument of the possibility of the unique sequence of PS108 arising from the misincorporation of nucleotides by reverse transcriptase did not provide a good explanation for the origin of the new sequence of PS108. On one hand, the conditions for cDNA synthesis was not the one that was conducive for AMV reverse transcriptase to misincorporate nucleotides as Mg^+, rather than Mn^+, is used and the concentrations of Mg^+ and nucleotides are not excessively high. On the other hand, the errors that the reverse transcriptase usually makes are known to be random point mutations by incorporating the wrong nucleotides in a sporadic fashion. Since the overall action of reverse transcriptase is dictated by the RNA and DNA templates (i.e. it will not generate a polynucleotide sequence without a template), it was unlikely that the 108-cassette was created by a nonspecific action of the reverse transcriptase. In addition, the replacement of the 57-cassette with a nonspecific sequence would have a reasonable probability of either shifting the reading frame or introducing a stop codon.

Taking all of this into consideration, it appeared that PS108 coded for a second species of the α subunit of pig heart SCS rather than being a cloning artifact. It still retained most of the sequence of PS57, including the active site histidine and the surrounding sequences.

In this context, it is worth mentioning that the presence of the structurally distinct protein isoforms in a regulated fashion correlating with cell differentiation, physiological conditions or different functions, is a characteristic phenomenon of several eukaryotic enzymes. For example, hemoglobins adopt several isoforms to suit the different physiological conditions in the embryonic, fetal and adult stages of development (Lewin, 1987, p400). The existence of multiple species of SCS in hearts has also been documented. As was discussed in the Chapter I, it had been argued that two kinds of SCS may be necessary in mammalian systems, one for the TCA cycle and the other for ketone body metabolism in the mammalian system; the first being ATP-specific, the other being GTP-specific. It has been demonstrated that there were indeed two kinds of SCS activity associating with two distinct proteins in mammalian tissues (Weitzman *et al.*, 1986; Weitzman, 1987). It has been our experience that a purified pig heart SCS preparation always shows a sharply defined band on SDS PAGE corresponding to the β subunit and a more diffuse band corresponding to the α subunit; in constrast, both α and β subunits of *E.coli* SCS yield only sharp bands on SDS PAGE (Figure 14). The purified pig heart SCS, if applied to a MonoQ™ FPLC column , can yield four peaks each of which shows similar activity (Figure 15). Moreover, an SCS with only ATP-specificity has been recently purified from pig heart in this laboratory by Mr. Edward Brownie.

It is not unreasonable to suppose that pig heart SCS may undergo an isoform change from fetus to infant considering the hemoglobin example as mentioned before, and that the PS108 could be a fetal remnant in the two day old piglet. Therefore, it is not difficult to rationalize that different kinds of mRNA's for SCS were found in the mammalian system.

Our current understanding of molecular genetics gives rise to two possible mechanisms for the generation of protein isoforms. The first of these is multiple gene systems. In this case, each protein isoform is encoded in a separate gene belonging to a multigene family (Lewin, 1987, p. 399). Each of the genes in the family is regulated in a different fashion, thus enabling specific isoforms to be expressed in different physiological conditions, development stages or differentiation states of the cell. The second possible mechanism is for a single gene system. Here, all the protein isoforms are encoded by a single gene and, through DNA rearrangement or alternative RNA splicing, the differential expression of multiple protein isoforms can be fulfilled (Breitbart *et al.*, 1987; Andreadis *et al.*, 1987).

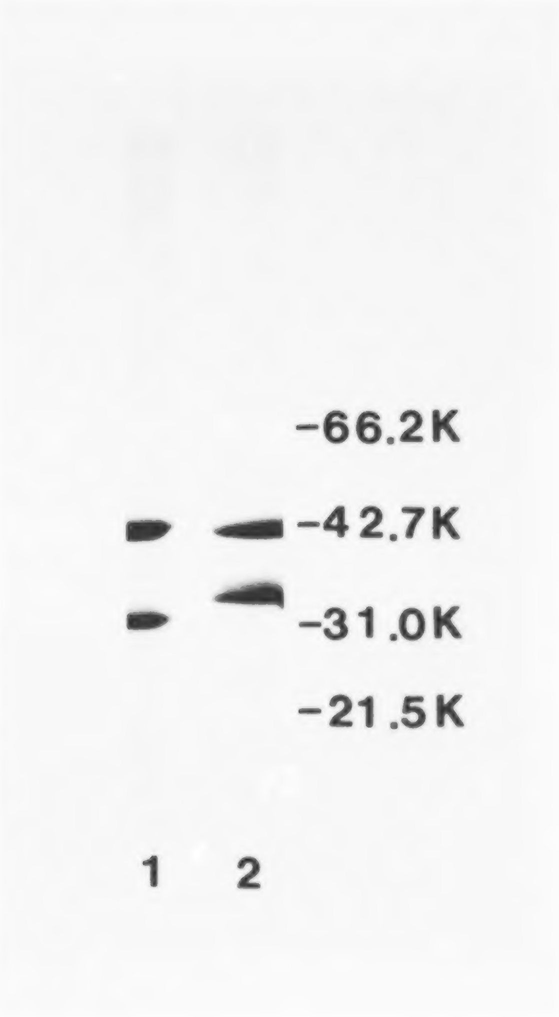


Figure 14. SDS PAGE analysis of *E.coli* and pig heart SCS The samples was analyzed in a 7 to 15% SDS PAGE (Material and methods). Sample in well 1 is *E.coli* SCS and sample in well 2 is pig heart SCS. The upper bands have been identifed as β subunits and the lower bands have been identified as α subunits.

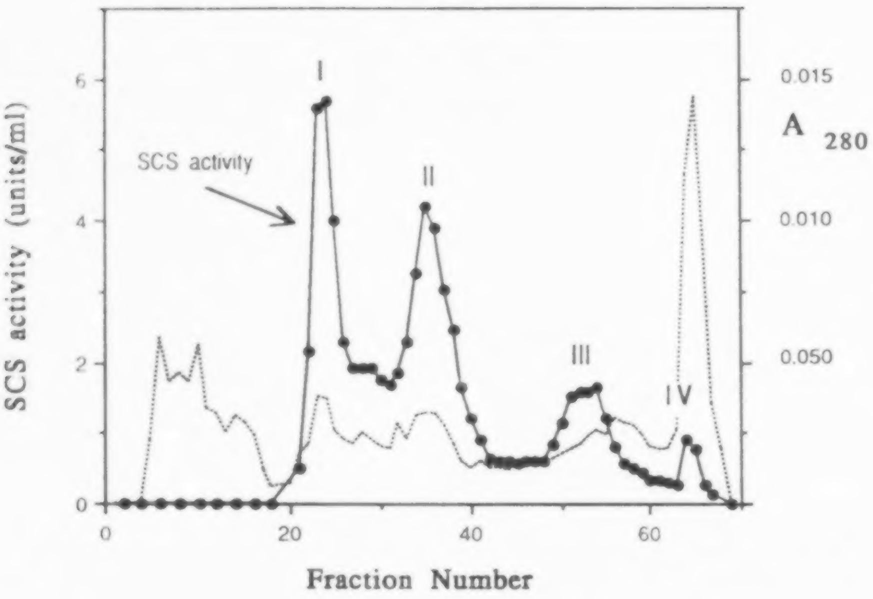


Figure 15. Separation of isozymes of pig heart SCS by FPLC
The SCS sample was applied to a MonoQ™ column in 50 mM Tris-HCl (pH8) and eluted with a NaCl gradient from 0 to 0.5 M. The SCS activity was measured by the standard spectrophotometric assay (Bridger *et al.*, 1969).

In the single gene system, the DNA rearrangement mechanism may be discounted because genetic replication normally occurs only during the course of DNA replication and subsequent cell division. Cells such as that of cardiac muscle are well differentiated and have reduced DNA replication activity. Therefore, it is argued that the isoforms of SCS are unlikely to be the products of DNA rearrangement. So far, the only protein isoforms found to be generated by the mechanism of DNA rearrangement are those relating to the immune system, such as immunoglobulins and T cell receptors (Hood *et al.*, 1985).

Both multiple genes and alternative RNA splicing of a single gene system thus represent plausible hypotheses to account for the generation of PS57 and PS108.

Several multiple gene systems have been documented and characterized; these include genes coding for globins (Collins and Weissman, 1984; Proudfoot, 1986), immunoglobulin (Hood, *et al.*, 1085; Hunkapiller and Hood, 1986), histones (Hentschel and Birnstiel, 1981; Old and Woodland, 1984), chorion proteins (Goldsmith and Kafatos, 1984), actins and tubulins (Firtel, 1981; Buckingham, 1985). Although the multiple gene system is a possibility for generating the two SCS isoforms, it is considered that PS57 and PS108 are better explained as the products of the single gene system and are generated via alternative RNA splicing for alternative splicing has been a common mechanism for producing protein isoforms in the well-differentiated cells of muscle and nerve (Andreadis *et al.*, 1987); and the fact that both PS57 and PS108 have the same non-coding sequences supports this hypothesis. If this is the case, the generation of PS57 and PS108 fits the mutually exclusive pattern of alternative RNA splicing, which, to date, has been detected only in genes for three kinds of contractile proteins: myosin light chain 1 and 3 of rat, mouse and chicken (Nabeshima, *et al.*, 1984; Periasamy, *et al.*, 1984; Robert, *et al.*, 1984; Strehler, *et al.*, 1985), α-tropomyosin of rat (Ruiz-Opazo, *et al.*, 1985; Ruiz-Opazo & Nadal-Ginard, 1987; Wieczorek, *et al.*, 1988), β-tropomyosin of rat (Helfman, *et al.*, 1986), and tropomyosin II of *Drosophila* (Karlik and Fyrberg, 1986), Troponin T of skeletal fast muscle of rat and quail (Breibart, *et al.*, 1985; Breibart & Nadal-Ginard, 1986; Hastings, *et al.*, 1985; Medford, *et al.*, 1984).

Above all the arguments, it was clear that more evidence would be required to confirm the authenticity of the PS108, and to identify the mechanism that produced it. The next chapter describes the cloning of pig genomic DNA, confirming alternative splicing as the origin of PS108.

CHAPTER V

GENOMIC DNA CLONING OF THE GENE FOR THE α SUBUNIT OF SCS

(1) INTRODUCTION

Genomic DNA cloning of the gene of the α subunit of SCS was the logical follow-up approach after the cDNA cloning and the discovery of a potential new isoform. This procedure could provide evidence for the authenticity of the 108-cassette if the sequence could be found to reside in the gene for the α subunit of SCS. Furthermore, it could provide a rationale for the origin of the two protein isoforms based on the knowledge of the organization of the genomic DNA encoding the α subunit of SCS. If the two different cDNA cassettes originated from alternative RNA splicing, the sequences of the two cassettes would be expected to be arranged in tandem in the genomic DNA; on the other hand, if the two cassettes were encoded in separate genes, each gene should only contain one of the cassettes.

The technological development of the polymerase chain reaction (Saiki *et al.*, 1988) appeared to offer a very attractive strategy to clone genomic DNA to determine the organization of the two cassettes. Oligonucleotides could be derived from the cDNA sequences in the constant regions flanking the cassettes and these could be used to amplify the specific fragment(s) of genomic DNA encoding the cassettes. A single band amplification pattern would support the alternative RNA splicing mechanism and a multiple band amplification pattern would be compatible with the possibility of a multiple gene system. There was one report in the literature concerning an investigation of intron structures by the means of PCR, which supported this approach (Bruzdzinski and

Gelehrter, 1989). However, after a lengthy series of investigations, it appeared that this strategy was not as promising as we had hoped.

Under carefully selected conditions, a single piece of DNA sized at 1.7kb (Figure 16) could be amplified from the pig genomic DNA by primers derived from the cDNA sequences flanking the cassettes; this seemed to support the mechanism of alternative RNA splicing. However, in hybridization experiments, this DNA fragment could hybridize only to a probe based on the 57-cassette at very low stringency even though it hybridized well with a probe based on the 108-cassette, thus raising some doubts about the authenticity of this DNA fragment. Worst of all, this DNA fragment was found difficult to subclone for further studies, despite extensive efforts. More PCR experiments aiming to obtain smaller fragments from the 1.7kb fragment using the different combinations of the oligonucleotides derived from the two cassettes and both the upstream and downstream sequences yielded little support for the authenticity of this DNA fragment.

PCR amplification experiments aimed at obtaining DNA fragments of the introns between the upstream sequence and either one of the cassettes or between either of the cassettes and the downstream sequence could, in some cases, yield unique DNA fragments. However, PCR experiments never led to a successful amplification of a DNA fragment that could be considered to be a good candidate for the putative intron between the two cassettes. All of these data produced no answer as to whether the two cassettes were encoded in two genes or if the target intron was simply too large to be amplified even though the two cassettes were arranged in tandem in one single gene.

It is known that introns can range in size from several hundred base pairs to larger than ten kilobase pairs. The normal upper limit of the size of a piece of DNA that can be amplified by PCR, however, is about 3kb; notwithstanding the report that a 10kb fragment of DNA was amplified which could be detected by autoradiography (Jeffreys *et al.*, 1988).

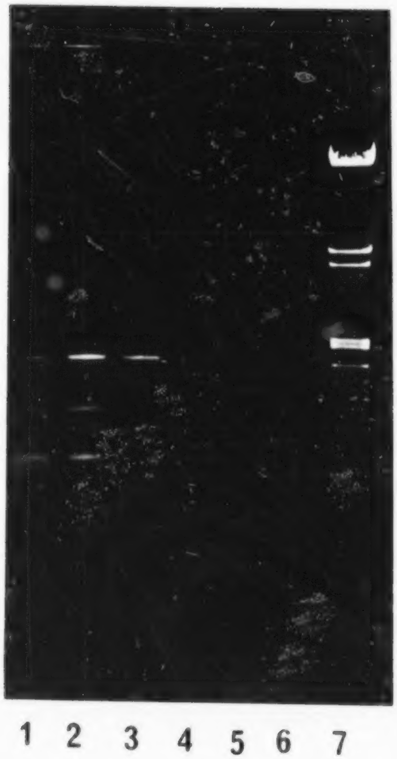


Figure 16. Amplification of a genomic DNA fragment This is the photograph of a 0.8% agarose gel. In well 7 was λ DNA digested with Hind III and EcoR I, used as the standards. In well 1 to 6, were PCR products using pig genomic DNA as the template. The primers were derived from the sequences immediate before and after the cassette in the cDNA. The amplifications were described by Saiki *et al.*, the annealing temperature was 68°C and the $MgCl_2$ concentrations were 4 mM (that in well 1) to 1.5 mM (well 6) with 0.5 mM difference between each reaction.

Therefore, we reached the conclusion that PCR alone might not be a sufficient tool to clone the desired gene fragment(s).

The logical solution to this problem was to apply the more traditional methods of cloning the gene(s) from genomic DNA libraries by hybridization experiments in the hope of isolating a genomic DNA clone having two cassettes arranged in tandem, or two clones each encoding a different gene for the α subunit of SCS.

(2) EXPERIMENTAL PROCEDURES

(a) Isolation of genomic DNA clones for the α subunit of SCS

cDNA clones of PS57 and PS108 were mixed and labeled by biotin-modified dUTP as described in Chapter IV. The probes thus obtained included all the sequence information available for the cDNA of the α subunits of SCS. The biotin-labelled cDNA was used to screen a pig genomic DNA library constructed in the vector EMBL3 (Material and Methods) and propagated in an *E. coli* strain of NM538 by procedures described in Chapter IV. When 1×10^5 clones were screened, two were found to be positive to the probes; these were designated as λE4 and λE6.

The λ DNA was isolated by basically the same procedures as described in Chapter IV with special attention paid to the multiplicity of infection (MOI) of the initial phage infection of the bacteria and the length of the time allowed for replication.

The MOI of 0.1 for the initial input of phages, as recommended (Frischauf, *et al.*, 1983; Frischauf, *et al.*, 1987), seemed to be important. However, due to the difficulty of obtaining 1×10^7 pfu (the normal input of bacteria was about 10^8 to 10^9) directly from the phage plaques, the desired titer was obtained with a gradual increment by using the higher titer of phage particles grown in each step to infect bacteria for a next step of phage growth.

The length of growing time also seemed crucial to produce enough phage particles for the isolation of λ DNA. Although different EMBL3 clones grow differently, a growing time of about 8 hours at 37 °C seemed to yield satisfactory titers and could allow synchronizing cell lysis.

(b) Oligonucleotide hybridizations

The oligonucleotides were end-labeled with ^{32}P for use as probes. In a typical experiment, 150 ng of oligonucleotides was mixed with 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA and 150 µCi of [γ-^{32}P]-ATP in a final volume of 50 µl. The reaction was initiated with the addition of 10 units of T4 DNA kinase and was allowed to proceed in 37°C for 45 to 60 minutes. The reaction was stopped by heating at 70°C for 10 minutes. The oligonucleotides thus labeled could be used directly in hybridization experiments. However, it was also a normal practice to separate the unincorporated [γ-^{32}P]-ATP and other nucleotides from the oligonucleotides in order to monitor the process of labeling, especially in the first experiment or when a component of the reaction was changed.

This separation of the oligonucleotides from the smaller molecules was achieved by ion exchange chromatography. The labeling mix was allowed to pass a total volume of about 200 µl of DEAE-Sephacel packed into a Pasteur pipet that was equilibrated with 10mM Tris-HCl (pH7.6) and 1mM EDTA. The resin was washed with a pipetful of the same buffer and the eluant was discarded. At this stage, essentially all the radioactivity was retained by the resin. A pipetful of 10 mM Tris-HCl (pH7.6), 1 mM EDTA and 0.2 M NaCl was then passed through the resin and brought with it the radioactivity associated with small molecules. Finally, the labeled oligonucleotides were eluted with about 0.5 ml (more than two volumes of the resin) of 10 mM Tris-HCl (pH 7.6), 1 mM EDTA and 0.5 M NaCl and then about 0.5 ml of 10 mM Tris-HCl (pH 7.6), 1 mM EDTA and 1 M NaCl. All the eluant was monitored by scintillation counting. In most experiments, more than 80% of the total radioactivity was found to be associated with the fractions containing oligonucleotides.

The filter replicas of the plaques produced on agar plates by the phages were obtained by basically the same methods described in Chapter VI. The filters were prehybridized overnight in 6xSSC, 0.2% BSA, 0.2% Ficoll, 0.2% polyvinylpyrrolidone and 50 µg/ml denatured sheared herring sperm DNA. The hybridization was carried out in the same solution plus the ^{32}P labeled oligonucleotide probes overnight (at 45°C in this case). Typically, 30 filters of 82mM diameter were hybridized in 60ml prehybridization solution supplemented with 150 ng of oligonucleotide probes in a 1 L beaker. The filters then were washed with 6xSSC and 0.1% SDS, first at room temperature and then twice at the hybridization temperature, 20 minutes each time, and finally at 5°C above the hybridization temperature for 20 minutes. The filters thus treated were rinsed in 6xSSC and subjected to autoradiography in -70°C with an intensifying screen.

(c) Polymerase chain reactions

The basic protocols of PCR were adopted (Saiki *et al.*, 1988) except that smaller amount of primers (about 250ng each) was used. The different parameters of the reactions were also changed from case to case to optimize the amplification of different target DNA's. It was our experience that the most important factors that determine the effectiveness of a PCR amplification were $MgCl_2$ concentration and annealing temperature. In typical trial experiments, several reactions were carried out in parallel in $MgCl_2$ concentrations ranging from 1.5mM to 3.5mM with 0.5mM difference in [$MgCl_2$] in each reaction mix. There were successful PCR experiments with $MgCl_2$ as low as 0.8mM.

DNA samples in various forms have been successfully applied in PCR as templates, which including DNA preparations of different degree of purity, λ phage, M13 phage and DNA embeded in both normal or low-melting point agarose gels.

For inverse PCR (Ochman, *et al.*, 1988; Silver and Keerikatte, 1989; Triglia, *et al.*, 1988), DNA samples were digested with restriction enzymes (enzymes that produced blunt or 5' overhanging ends were normally chosen) and then heated to stop the reactions before being diluted to less than 2 μg/ml of DNA in an extension and ligation mix of 50 mM Tris-HCl (pH7.5), 10 mM $MgCl_2$, 1.5 mM ATP and 0.2 mM of dATP, dGTP, dTTP and dCTP. The extension and ligation reactions were initiated by adding 6 units of Klenow fragments and 4 units of T4 DNA ligase and allowed to proceed overnight in room temperature. The ligated DNA was then precipitated and resuspended in TE buffer for the PCR.

(3) RESULTS

(a) Preliminary analysis of the two genomic DNA clones

To ensure that the two genomic DNA clones had the sequences of interest, hybridization experiments with two specific oligonucleotides were carried out. Oligo#1 represented a portion of the anti-mRNA sequence of the 57-cassette; its sequence was

5' TGAATATGGCCAGGCATGATG 3'.

Oligo#2 represented a portion of the anti-mRNA sequence of the 108-cassette; its sequence was

5' CTTCCTGTATTCTTCCTGCTG 3'.

Plaque hybridizations indicated that λE6 could hybridize only to Oligo#1 but not to Oligo#2; on the other hand, λE4 could hybridize to Oligo#2 but not to Oligo#1. Therefore, it could be concluded that each clone contained only one cassette — λE4 would contain only the108-cassette and λE6 only the 57-cassette.

Do these data support the multiple gene system hypothesis — or do they simply reflect different two regions in a single gene? To answer these questions, more hybridization experiments were carried out with the rationale that if λE4 and λE6 were two independent genes they might contain not only the cassette sequences but also the upstream and downstream sequences, since each EMBL3 clone should contain an insert sized from 10 to 22kb which could accommodate more than one intron.

Two more oligonucleotides were used. Oligo#3 represented a portion of the mRNA sense sequence of the exon upstream of the cassettes with the sequence

5' CACAGGCTGCTGCGCCAGGGAAA 3'

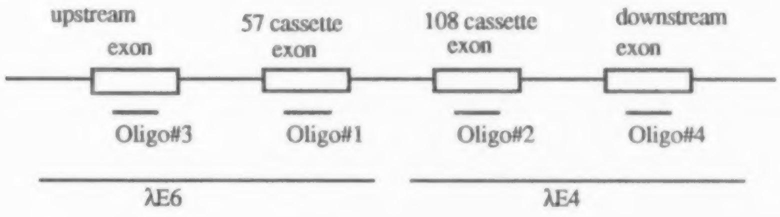
and Oligo#4 represented a portion of the anti-mRNA sense sequence in the exon downstream of the cassettes with the sequence

5' CCAACTTGTGTTGTTTCATGAACTGC 3'.

The results of the plaque hybridizations suggested that λE6, while perhaps containing the 57-cassette, could have the exon upstream of the cassettes since it could hybridize to Oligo#3, but not the exon downstream for it failed to hybridize to Oligo#4. In contrast, λE4 which might have the 108-cassette could have the downstream exon but not the upstream exon for it only hybridized to Oligo#4 but not to Oligo#3.

Summing up the oligonucleotide hybridization data, one cannot rule out the possibility of the two cassettes belonging to two separate genes, but the possibility of the two cassettes deriving from a single gene by alternative RNA splicing must be considered. This is the more logical explanation of the two cDNA molecules having identical coding and non-coding sequences outside the cassettes.

If the alternative RNA splicing is the right mechanism, the above data would indicate that the organization of the two cassettes in the genome would require the 57-cassette to precede the 108-cassette, as shown below:



Isolation and characterization of the λ DNA would help tremendously in the identification of these clones. After the isolation of the λ DNA and cleavage of the inserts from the vectors for analysis on agarose gels, it was found that λE4 had an insert of about 15kb and λE6 had an insert of about 14kb (Figure 17).

(b) The authenticity of the 108-cassette and the identification of the intron between the 108-cassette exon and the exon downstream

Unlike the 57-cassette, whose sequence had been found in many clones of the α subunit of SCS from diverse sources (*E. coli*, *Thermus aquaticus*, yeast, and rat liver), the 108-cassette of the clone of the α subunit of pig heart SCS was a previously unknown sequence and its authenticity was questioned ever since it was discovered. More doubts were cast on this matter because of the failure to amplify by PCR a segment of DNA from primers coding the upstream and downstream sequences of the cassettes. Therefore, it was considered to be important to obtain evidence of the genuineness of the 108-exon before planning any other work.

The isolation of λE4 and the demonstration that it might have the putative 108-cassette and the downstream intron (but not the 57-cassette) strongly implied an arrangement with the 108-cassette and the downstream exon being only one intron apart.

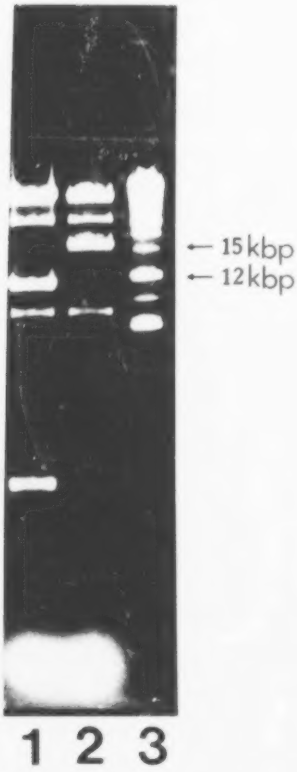


Figure 17. Sal I digestion of λE4 and λE6 The λ DNA digested with Sal I were analyzed by agarose gel (0.8%) electrophoresis. Lane1 is that from λE4 and the insert is split into two fragments of sizes about 3 and 11kbp. Lane 2 is that from λE6 and the insert is about 15kbp. Lane 3 is the high molecular weight marker from BRL, the 12 and 15kbp fragments are indicated by arrows.

This intron, along with parts of the 108-cassette and the downstream exon, might be able to be amplified by PCR for detailed studies to provide the first solid evidence of the authenticity of the 108-cassette.

The primers coding for the 108-cassette and the downstream exon were used for the PCR experiments and a DNA fragment of about 1.3kb in size was amplified using λE4 phage as template (Figure 18). This DNA possessed a Nde I restriction enzyme site, which could be correlated to a Nde I site in the 108-cassette. Encouraged by the restriction enzyme analysis data, this DNA fragment was promptly subcloned and subjected to DNA sequencing.

Figure 19 illustrates the DNA fragment derived from the sequencing experiments. At one end, after the primer sequence, we found the nucleotide sequence of the 108-cassette in the expected position, followed by a sequence that was not present in the cDNA, apparently belonging to an intron. Sequencing data obtained from the other end of this fragment gave the sequence following the cassettes in the cDNA.

This fragment of DNA was also later found to be part of a Bgl II restriction fragment of λE4, (as will be described later) ruling out the possibility of an artifact being generated by *T. aquaticus* DNA polymerase.

These data demonstrated that the cDNA containing the 108-cassette coded for an authentic isoform of SCS and that it was not an artifact of cDNA cloning in the sense that it was part of the gene of the α subunit of SCS. The data further suggested that the intron between the 108-cassette and the downstream exon was about 1.3kb in size.

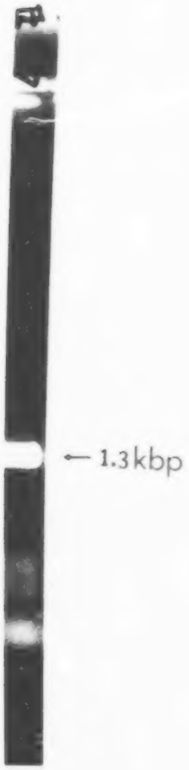


Figure 18. The DNA fragment containing the intron between 108-cassette and its downstream exon The fragment was amplifed according to Saiki *et al.* (1988) and analyzed in an 8% agarose gel.

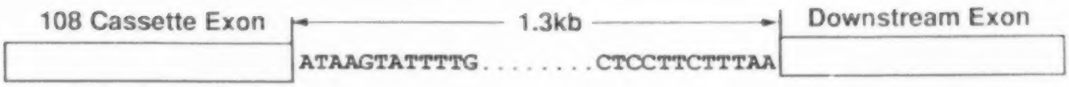

108 Cassette Exon
1.3kb
Downstream Exon
ATAAGTATTTTG........CTCCTTCTTTAA

Figure 19. The intron between the 108 cassette and the downstream exon

(c) The identification of the intron between the upstream sequence and the 57-cassette

PCR experiments using λE6 as templates and primers derived from the upstream sequence of the cassette and the sequence of the 57-cassette yielded a fragment of DNA of about 220bp (Figure 20), indicating the size of this particular intron to be relatively small, about 200bp. This piece of DNA was also subcloned and sequenced. The sequencing data confirmed that the amplified DNA fragment was indeed the one coding for the intron between the upstream exon and the 57-cassette exon (Figure 21).

(d) The intron between 57 and 108-cassette exons

The discovery of two partial clones of the gene of the α subunit of SCS, and the limited intron sizes between the upstream exon and the 57-cassette exon, 108-cassette exon and downstream exon suggested that it might be a long sequence between the 57-cassette and the downstream exon or between the upstream exon and the 108-cassette, since two random Mbo I digestions during the library construction occurred in the same region. The large size of the intron(s) could be the reason for the failure of the PCR experiments aimed at amplification of a fragment containing all three introns.

Efforts were made to isolate DNA fragments containing the sequences downstream of the 57-cassette and upstream of the 108-cassette to investigate the boundaries of the exons and the size of the intron.

Restriction enzymes known to have no site on the EMBL3 vector were chosen to analyze the clone λE6 in an attempt to isolate a fragment containing the 57-cassette exon and part of the intron at its 3' end (Figure 22). Among them, EcoR I was found to be able to split the larger insert fragment into two fragments, with sizes of about 7kb and 5kb. Agarose gel stubs containing these two DNA fragments were analyzed by PCR in order to

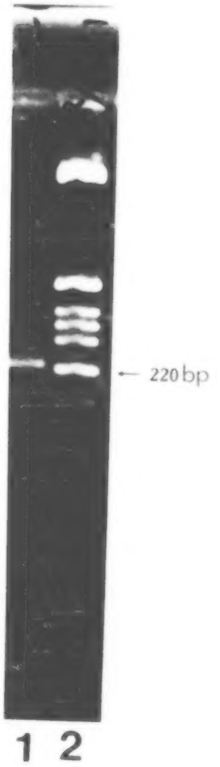


Figure 20. The DNA fragment residing the intron between 57-cassette and its downstream exon Lane 1 is the fragment amplified according to Saiki *et al.* (1988); lane 2 is the DNA fragments of pBR322 digested by Hinf I. The 220bp fragment is indicated by an arrow. The DNA was analyzed in a 3% low melting-point agarose gel (NuSeive™ from FMC).

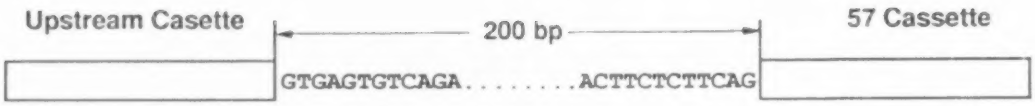

Upstream Casette
200 bp
57 Cassette
GTGAGTGTCAGA........ACTTCTCTTCAG

Figure 21. The intron between the upstream exon and the 57 cassette

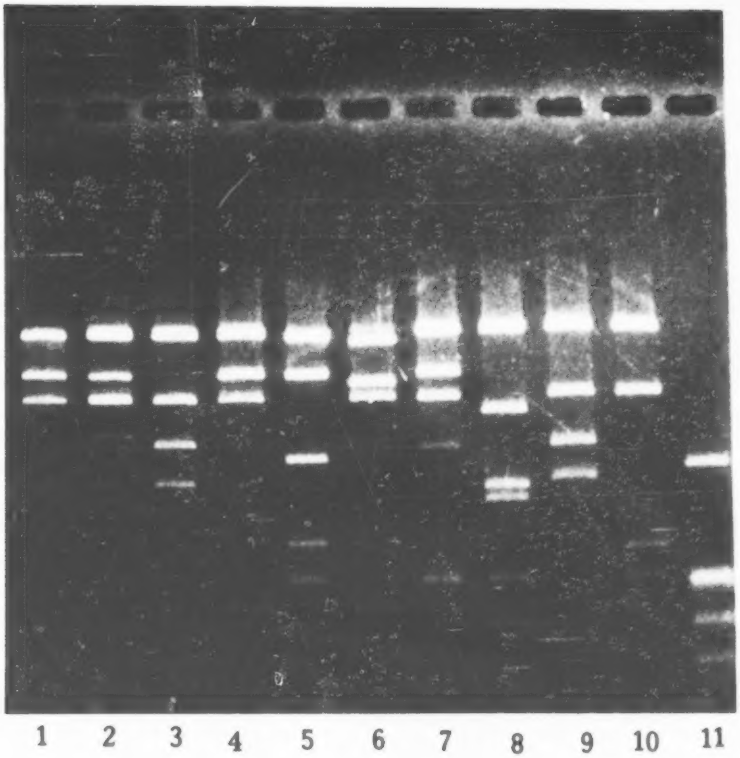
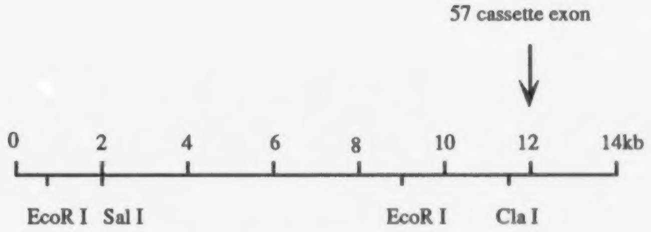


Figure 22 Restriction enzyme analysis of λE6 λE6 DNA was digested with Sal I and a second enzyme and then analyzed in a 0.8% agarose gel. The second enzyme was Xho I (lane 1); Xba I (lane 2); Sst I (lane 3); Spe I (lane 4); Sma I (lane 5); Sca I (lane 6); Nar I (lane 7); Hind III (lane 8); EcoR I (lane 9); BamH I (lane 10); Ava II (lane 11).
In lane 9, it is shown that the insert is split into two fragments with the sizes of 5 and 7kb, respectively.

determine which fragment containing the 57-cassette exon and surrounding introns. It was found that the PCR experiments used to generate the fragment containing the 57-cassette and its upstream exon could be repeated using the gel stub containing the 5kb fragment as templates and produced the expected pattern (Figure 20). Therefore, it was concluded that the 5kb fragment was the one harboring the 57-cassette, its upstream sequences and likely its downstream sequence. Subcloning and sequencing of this DNA fragment identified the boundary of the 57-cassette and its downstream intron. This was fully compatible with that of the downstream exon linked to the 108-cassette (Figure 23 and 24), as would be expected on the basis of linearity of the two cassette exons in the genomic DNA. Further restriction enzyme analysis showed that there was a Cla I site in this DNA fragment (Figure 25) and raised the possibility of locating this Cla I site to a position not too far away from the 57-cassette exon in the cDNA. If that was the case, deletion of the sequence preceding the 57-cassette from this Cla I site could help in revealing the length of the part of the intron in λE6. The deletion was made by the combined actions of EcoR I and Cla I and sequencing experiments were performed to confirm that this Cla I site was indeed correlated to the Cla I site in cDNA. Judging from the size of the fragment, it could be concluded that λE6 contained at least 2kb of the intron. Shown below is the schematic diagram of λE6:



The analysis of the intron borne by λE4 was done by a relatively straightforward and reliable method of "inverse PCR", to allow the investigation of the clones in a sequential fashion — "walking" along the DNA.

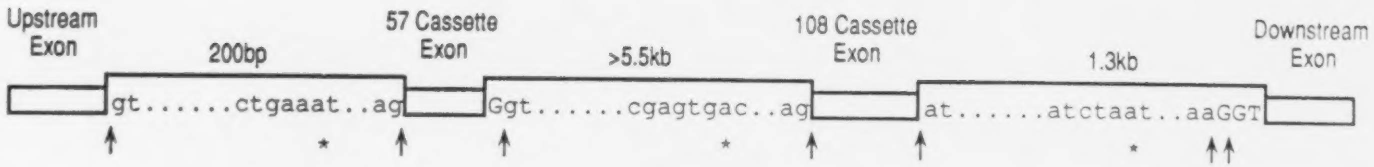

Upstream Exon
200bp
57 Cassette Exon
>5.5kb
108 Cassette Exon
1.3kb
Downstream Exon
gt......ctgaaat..ag
Ggt......cgagtgac..ag
at......atctaat..aaGGT
*
*
*

Figure 23. The organization of two mutually exclusive cassettes.

The arrows indicate the splicing points.
The * indicates the nucleotides which are suggested to form the branch points of putative lariats during intron excision.

Figure 24. Partial sequence of the genomic DNA of α subunit of SCS

```
ATCGATGCAGAAGTGCCCTTGGTCGTGTGCATCACTGAAGGCATCCCGCAGCAGGACATGGTGCGGGTAAAGCACAGGCTGCTG
CGCCAGGGAAAGACGAGGCTGATCGGGCCAAACTGCCCTGGAGTCATCAATgtgagtgtcagaaacaaagataagccttgattt
ttagcctctcaggggctctgagtcagtccaa.....tttgtgtgtgacctaggaggtataggtgtttttttttttttttttttt
tcacttctctgaaattgtgaattacttctcttcagCCTGGAGAATGCAAAATTGGCATCATGCCTGGCCATATTCACAAGAAAG
GACGAATTGgtgagtatgtcttttatagatttaaaaacaaaactggccctaaagctcaaggctcaaaactcatgacgagactct
cactgtgaatcaaatttattttaattaatttaactttgatgaaatttaccagggaaattttgattgtatatctgaatgacgttg
agaaacagtatacttctacagaatctagaagtatttttaggg.....agatctcccaggttttggtatgccggatgattttcag
ttgaaacctggatgtttggggtattatattatgaaattgggttttatgtgaaccttctgttttagctgactttaattctcactg
cttcagcagggatt.....ggcctcttagctctgtatctggatacacgaggcagaagcaaacccagggaacttaccgctgtgtc
actcttgagtctcgagtgactctgtcttctctctttcagAGATTTCTTAGATTTGCTGTATATATAACATCCAGGGTTTTTACT
TGTACTCAGCAGGAAGAATACAGGAAGATACCTCTGCTTTTTGGAACCAGAAGTATCCATATGataagtatttgaagtgatag
tctcttttactacgtatactttaagggtagcaagtataactgcccctaaaaacatttctttcaacactataaggttatgcagtt
tgcatttataaattaaatatcgtttacatagttagatgaaagccccttaaaagagtgacaaaatcctagtgttactatcccctg
attattatcctgatccatcaaaagattagaataaataaaatagtggtgcttcccagtaatctagggtgttacatgaacaagaac
ataattctattataaggatcaggtaggagtggaaaggagattacgtaattgcacacagagatgtacaactaagttaccgattat
ctttgtttaaaagctaatgcaccaaacgtacataggcattacctacctaatcttcttggttatgaaaggccccacagagtat..
...ggcaccactgacagtctcttattgttccactagaggagaacagtggtctgacttgaggaatcagtatgtgatctggagaca
gaagcctaactatatgaatttattttcctaaaccagggttattagccaaaaattaaaacttgaaatttatgctttttttcccngt
atggaccttttatagatttattattccatttactgtgtgtctgcagtagaaagtcattttatgtggaagacattaggtcttt
cttaactgctttcatagtcataaaacaactcaaactaataaaatacttccttctctatcttttaaaacctctctgttctaattt
aatccttttccctatttcattacctgacaatattgtttatttggatgtcagattatgaaatgggttttatatatagctgtgttt
gtaatttattttactgcatgataatcatatgtacacaatctaatatttttgaattctccttctttaaGGTATCGTGTCCAGATC
T
```

Legend: Combination of the partial sequences of two pig genomic DNA clones, based on the hypothesis that they are from a single gene. Upper case letters indicate exons and lower case letters indicate introns.

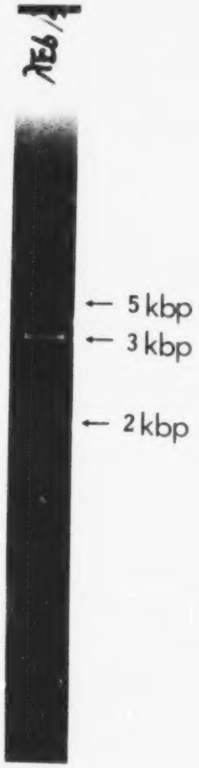


Figure 25 Cla I digestion of a DNA fragment from λE6 The 5kb fragment released from λE6 by the digestion of Sal I and EcoR I was isolated by glass powder method (Material and methods) and then digested by Cla I and analyzed in a 0.8% agarose gel. This is an incomplete digestion. The arrows show the undigested fragment (5kb) and two fragments (3 and 2kb) generated by Cla I digestion.

A limitation of conventional PCR is that it can only yield DNA fragments between two known sequences based on which oligonucleotide primers can be designed, but the regions that flank the two known sequences are excluded from analysis. Inverse PCR was designed to circumvent this problem by truncating the DNA by restriction enzymes and making a circle from the restricted DNA fragment before the amplification of the DNA by the extension of the primers in directions opposite to that of normal PCR. The products thus generated would be DNA fragments with primer sequences on both ends, sandwiching the sequence flanking the primers in the original sequence which could be subsequently analyzed. More primers could then be derived based on newly acquired sequences and more inverse PCR experiments could be carried out to extend the length of DNA being investigated. One major advantage of this method was that it could allow the walking of the DNA in the desired direction since restriction enzymes could be chosen as to eliminate the DNA extension in the undesired direction.

As mentioned above, there is an Nde I site in the end of 108-cassette exon; the action of this restriction enzyme would allow walking of the clone in the direction upstream of the 108-cassette exon in inverse PCR experiments.

A series of restriction enzymes paired with Nde I were chosen to digest the λE4 clone for the inverse PCR. It was found that a 900 bp fragment could be obtained by inverse PCR from the samples treated with Nde I and Bgl II (Figure 26), which implied that there was a Bgl II site 900 bp upstream of the 108-cassette exon. The knowledge of another Bgl II site 1.3kb downstream of the 108-cassette in the beginning of the downstream exon allowed the isolation of a 2.2kb Bgl II restriction fragment from the λE4 clone (Figure 26), which was used to derive the sequences at the upstream junction of the 108-cassette exon and 900 bp in the intron.

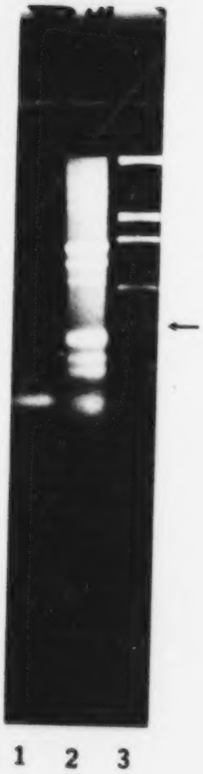


Figure 26. DNA fragments that contain the intron before 108-cassette This is a photograph of a 0.8% agarose gel. Lane 1, a 900bp fragment produced by inverse PCR. Lane 2, λ DNA digested with Hind III and EcoR I. Lane 3, DNA of λE4 digested with Bgl II; the 2.2kb fragment is indicated with the arrow.

The sequencing information on 2kb downstream of the 57-cassette exon and 900 bp upstream of the 108-cassette exon was used to design oligonucleotides for PCR experiments using pig genomic DNA as templates in an attempt to fill in the gap between the two known sequences. However, this attempt did not yield any DNA fragment, suggesting that there could still be a sizable sequence of the intron that was not being analyzed.

Analysis of the Bgl II restriction fragment also revealed that there was a Xho I site immediately prior to the 108-cassette. This Xho I site was determined to be a unique one in the insert of λE4, which could split the insert into two fragments of about a 3.5kb and a 11kb (Figure 26). These data suggested that there was at least 3.5kb of the intron residing in λE4.

(4) DISCUSSION AND CONCLUSION

Summarizing all the available sequencing information on both λE4 and λE6 (Figure 22 and 23), there are some features about the organization of the porcine genomic DNA that can be discussed.

A feature to be easily recognized was the size of the intron between 57 and 108-cassette exons, which would be larger than 5.5kb and much bigger than the one between the upstream exon and the 57-cassette exon. Could the difference of the sizes of the introns affect the its splicing and make PS57 the dominant species?

Examining the junctions of the exons revealed more interesting features. It is generally accepted that the DNA sequences of all introns in nuclear genes that code for proteins can be aligned by the GT...AG rule to show homology with a short consensus sequence at exon-intron boundaries

Exon...AG/GTAAGT...*Intron*...Py_{10}CAG/...*Exon*

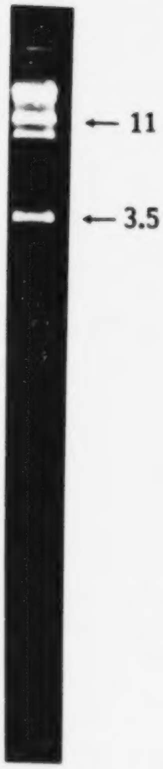


Figure 27. Digestion of λE4 with Sal I and Xho I The λE4 DNA was digested with Sal I and Xho I and analyzed in a 0.8% agarose gel. The arrows indicate the two fragments of inserts estimated at sizes 3.5 and 11kb. The other two fragments correspond to the two λ arms of the vector (EMBL3).

The dinucleotides of GT in the left junctions and AG at the right junctions of introns are believed to be always present (Lewin, 1987, pp473), despite exceptions can be found in the immunoglobulin λI J1-C1 gene (Sharp, 1981) and other genes (Jackson, 1991).

All the junctions in the three introns are in keeping with the GT...AG rule, except the one at the left junction and perhaps the one at the right junction of the intron between the 108-cassette and the downstream exon.

From the splicing pattern, it was obvious that the 57- and 108-cassette exons could not exist in the same cDNA and had to be "mutually excluded" from each other. The splicing of the intron between the 57-cassette and the downstream exon at the right junction would have the first nucleotide of a glycine codon retained with the 57-cassette. This requires joining the remaining nucleotides of the triplet retained in the downstream exon, not to the 108-cassette exon, to complete the glycine codon and maintain the reading frame. This arrangement makes this a *class I intron* (Sharp, 1981), which interrupts the codon between the first and second nucleotide. If the 57-cassette with an extra G resulted from the pattern of *class I intron* splicing had joined the 108-cassette, a stop codon would have been introduced and the translation of the protein would have terminated prematurely at the beginning of the 108-cassette.

The splicing at the left junction of the intron between the 108-cassette exon and the downstream exon certainly seemed unconventional. Although the junction had a sequence that could be aligned well with the consensus sequence, the crucial GT was changed to AT, this is generally believed to be unsuitable for splicing (Lewin, 1987, pp477). It is perhaps necessary to emphasize that the conclusion of the intron being bordered by AT at the left side is drawn from the solid sequencing data as shown in Figure 28. The discovery that AT could border the right junction of an intron thus challenges the existing dogma. Nevertheless, this AT junction seems to be less favorable for splicing and could severely

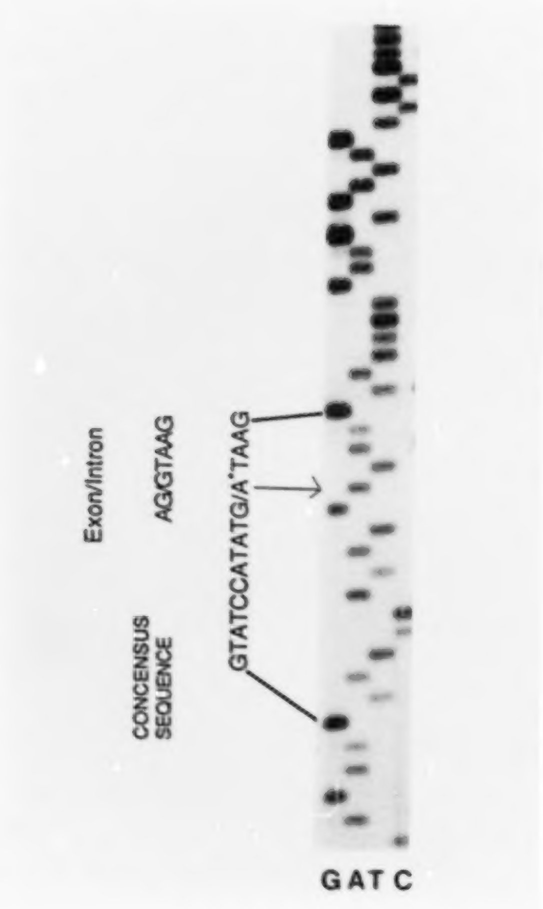

CONCENSUS SEQUENCE
Exon/Intron
AG/GTAAG
GTATCCATATG/A*TAAG
G A T C

Figure 28. The sequencing data of the boundary of 108-cassette
The sequencing reaction was performed by modified T7 DNA polymerase (Tabor and Richardson, 1989). The gel electrophoresis was done according to Hong (1987). The consensus sequence of intron/exon boundary is shown on the right. The arrow indicates the unambiguous A nucleotide.

limit the splicing of this intron. One could speculate that special factors might be required for the production of PS108 mRNA, perhaps providing an explanation for the fact that PS108 mRNA is a minor species compared to PS57 mRNA as was described in Chapter IV.

One may also speculate that the AT junction on the right side of an intron is non-splicable and is only introduced at certain development stage to nullify the 108-cassette to regulate the expression of SCS. However, this possibility can probably be ruled out, since it could only happen through a mechanism involving DNA rearrangement which would not be anticipated to occur in the well-differentiated hearts of piglets. The fact that PS108 mRNA can be found in the piglet heart indicates that the junction bordered by AT can be spliced properly, albeit at low frequency. Besides the our discovery, two examples of dinucleotide AT forming the right junctions of the introns have been found in genes coding for human proliferating cell nucleolar protein P20 and chicken cartilage matrix protein, although in a different type of consensus sequence (Jackson, 1991). Therefore, the categorical statements in textbooks regarding the conserved sequences at the intron junctions should perhaps be modified. Apparently, the lariats of the splicing intermediates could be formed not only by guanine nucleotides attaching to adenine nucleotides but by adenine nucleotides attaching to adenine nucleotides as well.

If the splicing pattern proposed at the left junction was correct, this splicing pattern could not produce an extra G to join the other two nucleotides to complete the reading frame in the same fashion as the alternative upstream intron. It is thus obvious that the right junction of the intron must be spliced differently for production of PS108 mRNA compared to PS57 mRNA, and that the splicing junction must be AA rather than AG at the right side of the junction to generate a complete glycine codon for the production of PS108. The intron between the 108-cassette exon and the downstream exon is a *class 0 intron* (Sharp, 1981), which interrupts the reading frame between codons. This "alternative" junction has not been reported before.

The putative TACTAAC boxes (Lewin, 1987, pp476) for the formation of lariats could be identified in all three introns (Figure 23).

It is worth emphasizing that, although it is still a possibility, there is no piece of evidence supporting the existence of a multiple gene family coding for the α subunit of SCS and the insert of λE6 being part of a pseudo-gene, except that the two cassettes were found in different partial genomic clones. Taking into the accounts that the PS57 and PS108 have the same non-coding sequences but the 57-cassette and 108-cassette have the different upstream, as well as downstream, introns, it is reasonable to suggest that the 57-cassette and 108-cassette do not belong to different genes of a multiple gene family. The apparent large introns following the 57-cassette and prior to 108-cassette provide the explanation for the fact that the two cassettes were found in two separate clones. Moreover, the compatibility of the 57-cassette with the downstream exon linked with the 108-cassette implies the linearity of 57- and 108-cassettes. There were preliminary plaque hybridization experiments indicating that the two genomic DNA clones λE4 and λE6 might actually overlap, which would support the alternative splicing mechanism. In order to yield more evidence to support the idea that an alternative splicing mechanism produces the two isoforms, it would be desirable to show that it is possible to generate a pig DNA fragment that can hybridize by Southern blot to two probes specific to the 57-cassette and 108-cassette. This experiment is being carried out by a fellow student with positive preliminary results.

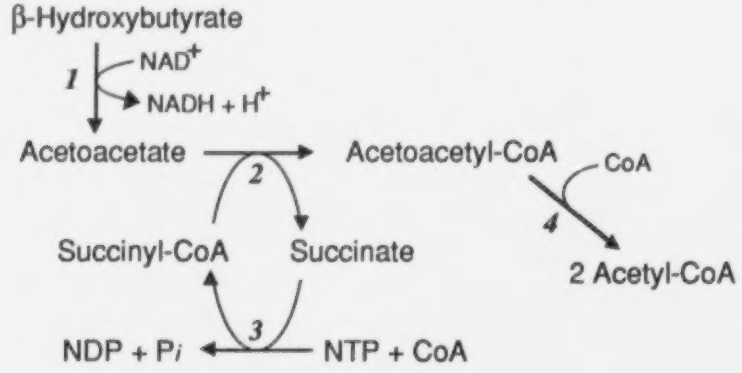
To summarize the results described in this chapter, we conclude that the genomic DNA cloning of the pig α subunit of SCS demonstrates the authenticity of the cDNA for the isozyme of pig heart PS108. It most likely originates through the mechanism of alternative RNA splicing. The demonstration of the corresponding protein variant and the identification of its function remain for future experimental analysis.

CHAPTER VI

SEQUENCE OF A cDNA CLONE ENCODING PIG HEART MITOCHONDRIAL CoA TRANSFERASE

(1) INTRODUCTION

Circulating ketone bodies (acetoacetate, β-hydroxybutyrate) provide fuel for many tissues under normal conditions, and the levels of these compounds rise markedly in certain pathological conditions such as diabetes or starvation. The route for the catabolism of these compounds in the mitochondria involves their conversion to acetyl-CoA by the pathway shown below:


β-Hydroxybutyrate
1
NAD+
NADH + H+
Acetoacetate
2
Acetoacetyl-CoA
CoA
4
Succinyl-CoA
Succinate
2 Acetyl-CoA
3
NDP + Pi
NTP + CoA

Route for catabolism of β-hydroxybutyrate and acetoacetate. The enzymes involved are: 1, β-hydroxybutyrate dehydrogenase; 2, CoA transferase; 3, succinyl-CoA synthetase; 4, acetoacetyl-CoA thiolase.

The subject of this chapter is the enzyme that catalyzes the transfer of the CoA moiety from succinate to acetoacetate, CoA transferase [3-ketoacid: succinyl-CoA Coenzyme A transferase (3-oxoacid CoA transferase, EC 2.8.3.5)].

As is mentioned in Chapter I, there is an extensive literature devoted to the study of the mechanism of action of CoA transferase, contributed in particular by Jencks and his colleagues. This group has shown that the overall reaction proceeds by a typical ping-pong mechanism that involves a covalent intermediate produced by thioester formation between a glutamate residue of the enzyme and CoA (Hersh and Jencks, 1976a, 1976b; Solomon and Jencks, 1969; Jencks, 1973). Formation of the E-CoA intermediate has been shown to proceed via an unstable anhydride species formed between the carboxylate groups of the enzyme and substrate (White and Jencks, 1976a; Pickart and Jencks, 1979). Elegant analysis of specific binding interactions (Moore and Jencks, 1982; Fierke and Jencks, 1986) has led to the proposal that intrinsic binding energy for the ADP moiety of the coenzyme is channeled to overcome energy barriers such as substrate destabilization.

In contrast to the close scrutiny that has been given to the mechanism of catalysis by CoA transferase, comparatively little attention has been given to the structure of the enzyme. By sedimentation analyses, the molecular weights of pig heart (White and Jencks, 1976b) and sheep kidney (Sharp and Edwards, 1978) CoA transferases were estimated to be 92,000 and 102,000, respectively. The M_r in the presence of denaturants was shown to be approximately 50,000, indicating the enzyme to be a dimer. It has been noted that both the sheep kidney (Sharp and Edwards, 1978) and pig heart (Moore and Jencks, 1982) enzymes can give rise to smaller fragments with M_r corresponding to about 29,000 and 25,000, presumably by proteolytic digestion during purification. The nicked enzyme retains full catalytic activity, suggesting that cleavage may take place at a susceptible site between domains.

In this chapter, the isolation and sequence of a cDNA clone encoding the cytoplasmic precursor to mitochondrial pig heart CoA transferase are described. The unusual characteristics of the amino acid sequence near the cleavage site are also discussed.

(2) MATERIALS AND METHODS

(a) cDNA cloning of the CoA transferase

The screening of the λgt11 library was done essentially by the methods described in Chapter V, except the washing of the filters in the hybridization experiments was done twice at 45°C (which was the hybridization temperature) in 3xSSC, 0.5% SDS for 20 minutes once at 50°C once for 20 minutes in the same solution.

The preparation of λ DNA and the sequencing of the target insert were done as described in Chapter IV.

(b) Purification of pig heart CoA transferase

The procedure used was an adaptation of that of Moore and Jencks (1982). DEAE-Sephacryl was substituted for DEAE-cellulose for the ion exchange chromatography step, and Affigel Blue™ (BioRad) was used for the final affinity chromatography step. These CoA transferase preparations were homogeneous by SDS-PAGE (see Figure 29) with specific activities of 13 to 16 Stern units/ng. We found that the presence of 0.1 mM phenylmethane sulfonyl fluoride (PMSF) and 0.1 mM EDTA in all buffers protected the enzyme from the previously noted proteolytic nicking during preparation (also see Results and Discussion). Omission of PMSF from the buffer, even for the final affinity chromatography step, yielded a product that was nicked (Figure 28) but with full specific activity.

(c) Protein sequence determination

Protein sequence information was required both in order to construct oligonucleotide probes and to confirm the identity of any cloned cDNAs (see Results and Discussion). We took advantage of the aforementioned proteolytic nicking to obtain

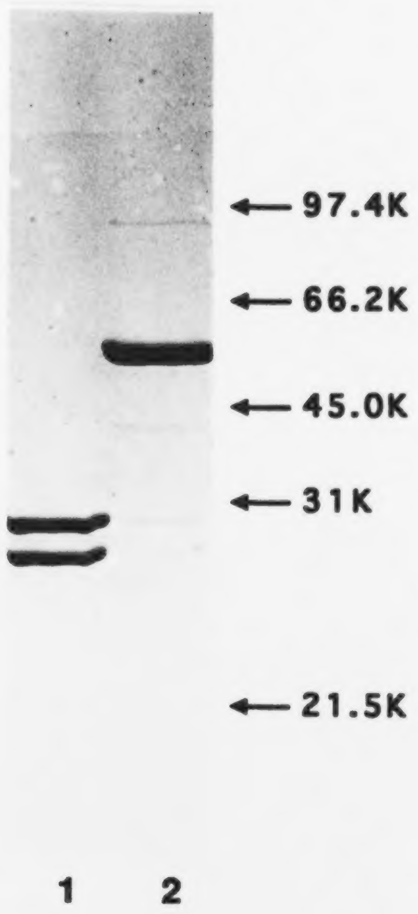


Figure 29. SDS PAGE analysis of nicked and unnicked prepatations of CoA transferase The electrophoresis was performed using a 10% SDS PAGE. Lane 1, nicked enzyme, prepared by omission of protease inhibitors from final purification step. Lane 2, unnicked mature enzyme

sequence information from two regions of the polypeptide. The two proteolytic fragments were separated by SDS-PAGE (see Figure 28) and samples for sequence determination were prepared by electrophoretic elution as described by Hunkapiller *et al.* (1983). Protein sequence determination was performed by automated Edman degradation using an Applied Biosystems Model 470A instrument.

(3) RESULTS AND DISCUSSION

(a) N-Terminal sequence determination

Sequential Edman degradation of mature CoA transferase and its proteolytic fragments revealed that the enzyme and its larger (M_r = 29K) fragment had the N-terminal sequence:

Thr-Lys-Phe-Tyr-Thr-Asp-Ala-Val-Glu-Ala-Val-Lys-Asp-Ile-Pro-Asn-Gly-Ala-
Thr-Val-Leu-Val-Gly-Gly-Phe-(Gly or Arg)-Leu-(Ser or Cys or Arg)-Gly-Ile-Pro-.

The N-terminal sequence of the smaller (M_r = 25K) fragment was determined to be:

Lys-Leu-Gly-Asp-Asn-Val-Arg(?)-Glu-Arg(?)-Ile-.

(b) cDNA clone isolation and sequencing

The sequence information was used to design oligonucleotide probes using established codon usage guidelines (Lathe, 1985). Two probes were based upon the sequence of the larger (M_r = 29K) fragment. The mRNA-sense Probe I was derived from Thr^1 to Lys^{12} as:

5'ACCAGGTTCTACACAGATGCTGTGGAGGCTGTGAA3'.

The mRNA-sense Probe II was derived from Ile^{14} to Phe^{25} as:

5'ATCCCCAATGGCGCCACAGTGCTGGTGGGCGGCTT3'.

A screening of about 1.5×10^5 pfu of the pig heart cDNA library yielded 3 clones that were found to hybridize to both of these probes. One of these, designated λT6, was selected for further analysis. The λ DNA was purified and the insert was excised by treatment with EcoRI, resulting in two fragments (~ 1.6 kbp and 300 bp) separated on agarose gel and indicating the presence of an internal EcoRI site. The sequences of these fragments were determined according to the strategy outlined in Figure 30, and the complete sequence is shown on Figure 31.

(c) CoA transferase sequence

A single open reading frame encoding a protein of 520 amino acid residues was derived from the DNA sequence of λT6 (see Figure 31). Confirmation of this sequence as that of CoA transferase was provided by the identification of the amino acid sequences established by automated Edman degradation of the proteolytic fragments of the protein (see boxed sequences, Figure 31). The arrangement of these fragments clearly establishes that the larger proteolytic fragment is derived from the N-terminal portion of the mature protein, and that the smaller fragment is from the C-terminal portion. Residues 1 to 39, upstream from the known N-terminal sequence of the mature protein, constitute a signal sequence with typical characteristics for targeting to the mitochondrial matrix (von Heijne, 1986), including an abundance of arginine, serine, and small aliphatic residues, absence of acidic residues, and potential for formation of an amphiphilic α-helix. The sizes of the various portions of CoA transferase are summarized in Table I. The amino acid composition of mature CoA transferase as measured by amino acid analysis of the purified enzyme and as deduced from the derived sequence are shown in Table II. The excellent agreement confirms the identity of the cDNA clone that we have isolated.

Scanning of all available data bases failed to detect any proteins having significant homology to CoA transferase. Significantly, no regions of sequence homology could be found between CoA transferase and either the β-subunit of succinyl-CoA synthetase

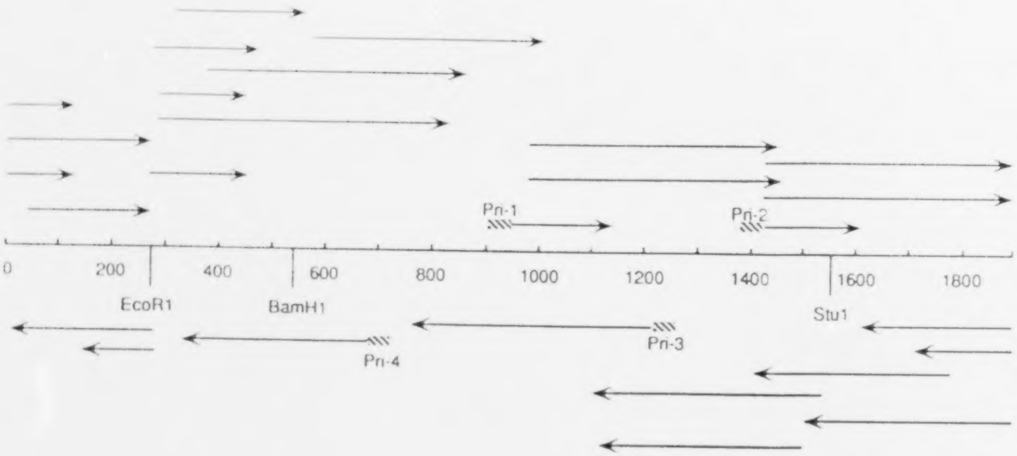

Pri-1
Pri-2
0
200
400
600
800
1000
1200
1400
1600
1800
EcoR1
BamH1
Stu1
Pri-3
Pri-4

Figure 30. Strategy used for sequence determination of λT6
Each arrow represents two sequencing reactions, one using dITP and the other using dGTP. The hatched bars labeled *pri-1* to *pri-4* indicate sites of usage of synthetic oligonucleotides as primers.

```
1                                                                                81
GAATTCCGACCGGTCCCGGCGGAGCAGTTGACACACTGAGAGGCGCTCCAACCCCGGCCTGCGCCCGCCGCCCTCCCGACC

ATGGCGGCTCTCACACTCCTCTCCTCCCGCCTGCGCCTCTGCGCCTCCGCCTACCGCTCCGGGGGCGCCTGGTCCCAGGGATGTGCTGGC
MetAlaAlaLeuThrLeuLeuSerSerArgLeuArgLeuCysAlaSerAlaTyrArgSerGlyGlyAlaTrpSerGlnGlyCysAlaGly 30

TACTTTTCAACTAGCACGCGTCGTCATACCAAATTTTATACGGATGCAGTGGAAGCTGTAAAAGACATCCCTAATGGTGCAACAGTGCTG
TyrPheSerThrSerThrArgArgHisThrLysPheTyrThrAspAlaValGluAlaValLysAspIleProAsnGlyAlaThrValLeu 60

GTTGGTGGTTTTGGGTTATGTGGAATTCCGGAGAATCTTATAGGAGCTTTGCTAAAGACTGGAGTAAAAGAACTTACTGCAGTCAGCAAC
ValGlyGlyPheGlyLeuCysGlyIleProGluAsnLeuIleGlyAlaLeuLeuLysThrGlyValLysGluLeuThrAlaValSerAsn 90

AATGCAGGGGTTGACAACTTTGGTTTGGGCCTTTTACTTCAATCCAAGCAGATAAAACGCATGATCTCTTCATACGTGGGAGAAAATGCA
AsnAlaGlyValAspAsnPheGlyLeuGlyLeuLeuLeuGlnSerLysGlnIleLysArgMetIleSerSerTyrValGlyGluAsnAla 120

GAATTTGAACGACAGTACTTAGCTGGTGAATTAGAAGTAGAGCTGACACCTCAGGGCACCCTGGCAGAGAGGATCCGTGCAGGTGGGGCA
GluPheGluArgGlnTyrLeuAlaGlyGluLeuGluValGluLeuThrProGlnGlyThrLeuAlaGluArgIleArgAlaGlyGlyAla 150

GGTGTTCCTGCGTTTTACACCAGCACAGGGTATGGCACCCTGGTGCAAGAAGGAGGGTCCCCGATCAAATACAACAAAGACGGCAGCATT
GlyValProAlaPheTyrThrSerThrGlyTyrGlyThrLeuValGlnGluGlyGlySerProIleLysTyrAsnLysAspGlySerIle 180

GCCATTGCCAGTAAACCGAGAGAGGTGAGGGAATTTAATGGTCAGCACTTTATTTTGGAGGAAGCAATCAGAGGGGATTTTGCTTTGGTG
AlaIleAlaSerLysProArgGluValArgGluPheAsnGlyGlnHisPheIleLeuGluGluAlaIleArgGlyAspPheAlaLeuVal 210

AAAGCCTGGAAGGCAGACCAAGCTGGAAACGTGACTTTCAGGAAAAGTGCAAGAAATTTCAACCTGCCGATGTGCAAAGCTGCAGAAACC
LysAlaTrpLysAlaAspGlnAlaGlyAsnValThrPheArgLysSerAlaArgAsnPheAsnLeuProMetCysLysAlaAlaGluThr 240

ACAGTGGTGGAGGTTGAAGAAATTGTGGATATTGGATCATTTGCCCCAGAGGACATTCATATTCCTAAGATTTATGTACATCGCCTTGTA
ThrValValGluValGluGluIleValAspIleGlySerPheAlaProGluAspIleHisIleProLysIleTyrValHisArgLeuVal 270

AAGGGAGAAAAATATGAGAAAAGAATTGAGCGTTTATCAGTCCGGAAAGAGGAAGATGTAAAAACCAGATCTGGTAAACTTGGAGATAAT
LysGlyGluLysTyrGluLysArgIleGluArgLeuSerValArgLysGluGluAspValLysThrArgSerGlyLysLeuGlyAspAsn 300

GTAAGAGAACGGATCATCAAGCGGGCGGCTCTTGAATTTGAGGATGGCATGTATGCTAATTTGGGCATAGGAATCCCACTTCTGGCCAGC
ValArgGluArgIleIleLysArgAlaAlaLeuGluPheGluAspGlyMetTyrAlaAsnLeuGlyIleGlyIleProLeuLeuAlaSer 330

AACTTTATCAGCCCAAATATGACTGTCCACCTGCAAAGTGAAAATGGAATCCTGGGTTTGGGTCCATATCCTTTACAAAATGAAGTTGAT
AsnPheIleSerProAsnMetThrValHisLeuGlnSerGluAsnGlyIleLeuGlyLeuGlyProTyrProLeuGlnAsnGluValAsp 360

GCAGATCTAATCAATGCAGGCAAGGAAACAGTTACTGTTCTTCCAGGAGCCTCTTACTTCTCCAGCGATGAATCATTTGCAATGATTAGA
AlaAspLeuIleAsnAlaGlyLysGluThrValThrValLeuProGlyAlaSerTyrPheSerSerAspGluSerPheAlaMetIleArg 390

GGGGGACATGTCAATCTAACAATGCTAGGAGCCATGCAGGTTTCCAAATATGGTGACCTGGCTAACTGGATGATACCTGGGAAGCTGGTG
GlyGlyHisValAsnLeuThrMetLeuGlyAlaMetGlnValSerLysTyrGlyAspLeuAlaAsnTrpMetIleProGlyLysLeuVal 420

AAAGGAATGGGAGGGGCTATGGATCTAGTGTCCAGTGCCAAAACCAAAGTGGTGGTCACCATGGAGCATTCTGCAAAGGGAAATGCACAT
LysGlyMetGlyGlyAlaMetAspLeuValSerSerAlaLysThrLysValValValThrMetGluHisSerAlaLysGlyAsnAlaHis 450

AAAATCATGGAGAAATGTACATTACCACTGACGGGAAAGCAATGTGTCAACCGCATCATTACTGAAAAGGCAGTGTTTGATGTGGACAGA
LysIleMetGluLysCysThrLeuProLeuThrGlyLysGlnCysValAsnArgIleIleThrGluLysAlaValPheAspValAspArg 480

AAGAAAGGTCTGACTCTGATTGAACTCTGGGAAGGCCTGACAGTGGATGACATAAAAAAGAGCACTGGGTGTGATTTTGCTGTTTCACCA
LysLysGlyLeuThrLeuIleGluLeuTrpGluGlyLeuThrValAspAspIleLysLysSerThrGlyCysAspPheAlaValSerPro 510

AAACTGATACCAATGCAGCAGGTCACAACT
LysLeuIleProMetGlnGlnValThrThr 520
1642
TGAAATGTGGAACAGATATTTGTCCCAAGCTGTGTGCTTTTCATTTTAGCCACATAGGATTTTATTGAAAAGACATCAGTAATTGTAATT

GTATTCAAACAAGTAGTTTTTGACTCGTTTTCTTCTAGTGTCTTATGCTTTATACAGCCATATAGACTTTGTTCTCTCAGATGTACTCTG

                             1851
ACATTTCAAGGAAAAACAAAAAAAAAAAGG
```

Figure 31. Sequence of λT6 and deduced amino acid sequence of CoA transferase Numbers in italics correspond to the protein sequence, numbers in plain text refer to the nucleotide sequence. Amino acid residues 1-30 (dotted underline) indicate the putative signal sequence for targeting to the mitochondrial matrix. Boxed residues (40-80 and 295-305) indicate the sequences determined by automatic Edman degradation analysis of the large and small proteolytic fragments, respectively.

(Buck, *et al*, 1985)[1] or acetoacetyl-CoA thiolase (Fukao, *et al*, 1989), two enzymes that interact specifically with succinyl- and acetoacetyl-CoA, respectively.

The sequence of CoA transferase provides a clear explanation for the marked susceptibility of the enzyme to proteolytic cleavage. Since the nicked enzyme has full catalytic activity, one might suppose a structure having the polypeptide chain folded into two distinct interacting domains, joined by a solvent-accessible bridge or hinge region, perhaps similar to the two-domain structure of phosphoglycerate kinase (Banks, *et al*, 1979). This concept is supported by the hydrophilicity plot for CoA transferase (Figure 31), showing that the site of proteolytic nicking (at residue 295) is located in an extremely hydrophilic region of the molecule. Further examination of the sequence upstream of the cleavage site shows that 15 of the preceding 25 residues are charged, clearly consistent with a surface location. We have noted a periodicity in charge distribution within this region; plotting residues 271 to 288 on a helical wheel (Figure 32) suggests that the bridge region could form an α-helix with positive and negative charges segregated on opposing surfaces.

Although it has been established that the active site contains a glutamate residue that is esterified to CoA during the course of the reaction (Solomon and Jencks, 1969), there is no information in the literature to identify one of the 37 glutamates as a candidate. In this connection, however, it has been shown that CoA transferase can undergo a mechanism-based fragmentation that is believed to occur at the thioesterified residue (Howard *et al.*, 1986). Although the precise site of cleavage has not been determined and its approximate

---

[1] The cited reference is to *E. coli* succinyl-CoA synthetase. In addition, no homology could be found between CoA transferase and the β subunit of pig heart succinyl-CoA synthetase (Bailey, D.L and Bridger, W.A., in preparation).

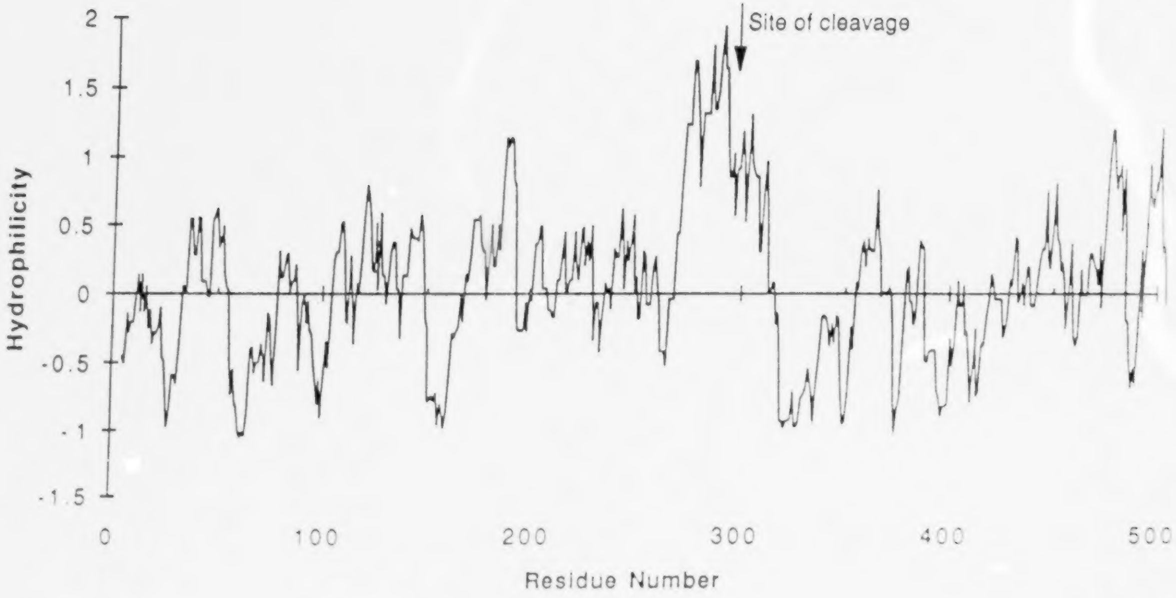


Figure 32. Hydrophilicity plot of the amino acid sequence of CoA transferase
The data were calculated using the program DNA Inspector™, using 8-residue window.

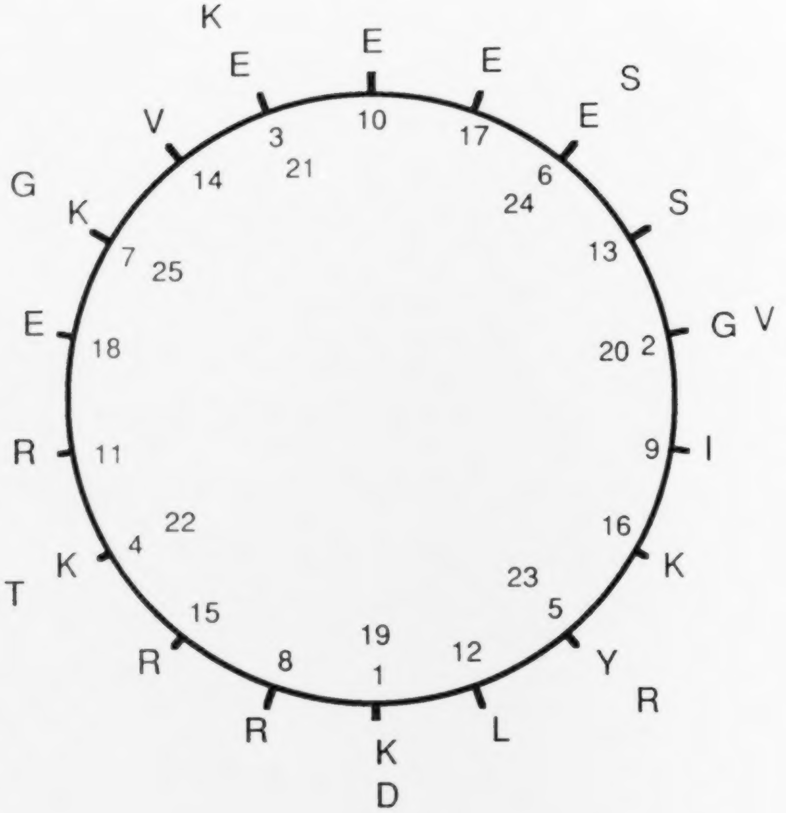

K
E
E
E
S
E
V
3
10
17
14
21
6
G
24
K
S
13
7
25
E
G V
20 2
18
R
11
9
I
22
16
4
K
K
23
T
15
5
19
R
12
Y
8
1
R
R
K
L
D

Figure 33. Helical wheel plot of CoA Transferase residues 271-295. This is the region of the molecule immediately preceding the site of proteolytic cleavage. The plot shows that if an α-helical secondary structure is assumed, the positively and negatively charged amino acid side chains would be segregated on opposite sides of the helix.

location could implicate *E*344, *E*358, and perhaps *E*369 as possible candidates for the site of thioesterification. It is hoped that this work might encourage further characterization of the active site region to identify its location within the sequence.

TABLE I

Sizes of various portions of pig heart CoA transferase

Sizes are calculated from deduced amino acid sequence (see Figure 19)

	Number of amino acid residues	Molecular weight
Precursor (with signal sequence)	520	56,371
Mature enzyme	481	52,197
Larger proteolytic fragment	256	27,906
Smaller proteolytic fragment	225	24,308

TABLE II

Amino acid composition of CoA transferase

Amino acid analysis was performed on the purified form of pig heart CoA transferase according to standard methods. The composition shown is based on a molecular weight of 52,200, and is compared to that derived from the composition based upon the sequence deduced from the cDNA.

Amino acid	Composition from amino acid analysis	Composition from cDNA sequence
Ala	39.3	39
Arg	19.3	20
Asx	43.0	43
Cys	N.D.	5
Glx	50.6	50
Gly	48.5	49
His	7.7	7
Ile	31.5	32
Leu	44.8	43
Lys	36.8	37
Met	15.0	13
Phe	17.1	17
Pro	18.8	18
Ser	25.8	25
Thr	26.8	28
Trp	N.D.	3
Tyr	12.0	12
Val	40.3	40

N.D.: not determined

CONCLUSIONS AND PERSPECTIVE

The main goals of the projects of cloning the cDNA coding for the α subunit of SCS and CoA transferase have been achieved. Moreover, the discovery of PS108 provides the strong evidence of the existence of multiple species of SCS isoforms and could offer a possible rationale for some of the different functions of SCS. The genomic DNA cloning has helped us to obtain evidence about the authenticity of a newly discovered clone of PS108 and to understand its origin.

There are many questions that still need to be addressed. For example, we have no information about in which stage of the cell development or in which tissues the isoform coded by PS108 is most active. This question may be answered by Western blots and Northern blots of samples from different tissues of pigs of various ages or stages of development, and probing these with antibodies against the specific peptides coded by 108 or 57-cassettes and/or with oligonucleotides derived from these cassettes. Specific antibodies may also prove useful for the identification of the locations of different SCS isoforms in mitochondria. However, there should be a cautious note that the synthetic peptides are not always good antigens.

The function of the protein sequence coded by the 108-cassette, which one might speculate could be crucial for SCS to act in either TCA cycle or ketone body metabolism, could also be investigated. This might be done by introducing specific restriction sites into both the PS108 clone and the SucD gene of *E. coli* and, subsequently, producing a chimeric SCS with *E. coli* SCS as the main skeleton bearing the protein sequence coded by 108-cassette. This would take advantage of the system for overexpression of the SucCD genes and the function of 108-cassette coded sequence could be tested directly. Probes specific for the 108-cassette should also be used to screen other mammalian cDNA libraries to clone the PS108 counterpart in other species, if it exists.

It is likely that there are other species of SCS that remain to be identified, and it is obviously a possibility that the large intron between the 57 and 108-cassettes could accommodate other cassettes. Screening of the pig heart cDNA library by probes derived from this intron is desirable and potentially profitable.

A study of the pattern of the alternative splicing to produce either PS57 or PS108 could also be an interesting project. Cloning of the full gene of SCS and subsequently expressing the gene in different cell lines could set the stage for such an investigation. The mutation at the junction sequences or the switch of the locations of the two cassettes could all be fruitful.

The availability of the derived protein sequence of CoA transferase should be a big impetus for the determination of its 3-dimensional structure of this enzyme by x-ray crystallography, especially since the enzyme can be prepared in large quantities. The overexpression of CoA transferase would also be a natural follow-up project for the research in this field.

REFERENCES

Andreadis, A., Gallego, M. E. and Nadal-Ginard, B. (1987) *Ann. Rev. Cell Biol.*, 3: 207

Ansorge, W. and Barker, R. (1984) *J. Biochem. Biophys. Method.*, 9: 33

Ansorge, W. and De Maeyer, L. (1980) *J. Chromatogr.*, 202: 45

Ansorge, W. and Labeit, S. (1984) *J. Biochem. Biophys. Method.*, 10: 237

Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M. and Phillips, A. W. (1979) *Nature*, 279: 773

Berger, S. L. and Kimmel, A. R. (1987) *Method. Enzymol.*, 152

Breitbart, R. E., Andreadis, A. and Nadal-Ginard, B. (1987) *Ann. Rev. Biochem.*, 56: 467

Breitbart, R. E. and Nadal-Ginard, B. (1986) *J. Mol. Biol.*, 188: 313

Breitbart, R. E., Nguyen, H., Medford, R., Destree, A., Mahdavi, V. and Nadal-Ginard, B. (1985) *Cell*, 41: 67

Bridger, W. A. (1971) *Biochem. Biophys. Res. Comm.*, 42: 948

Bridger, W. A. (1974) *The Enzymes* (Boyer, P. D., ed) 3rd ed. X, pp581 Academic Press

Bridger, W. A., Ramaley, R. F., and Boyer, P. D. (1969) *Method. Enzymol.*, XIII: 70

Brownie, E. R. and Bridger, W. A. (1972) *Can. J. Biochem*, 50: 719

Bruzdzinski, C. J. and Gelehrter, T. D. (1989) *DNA*, 8: 691

Buck, D. and Guest, J. R. (1989) *Biochem. J.* 260: 737

Buck, D., Spencer, M. E. and Guest, J. R. (1985) *Biochem*, 24: 6245

Buck, D., Spencer, M. E. and Guest, J. R. (1986) *J. Gen. Microbiol.*, 132: 1753

Buckingham, M. E. (1985) *Essay Biochem.*, 20: 77

Collins, F. S. and Weissman, S. M. (1984) *Prog. Nucl. Acid. Res. Mol. Biol.*, 31: 315

Davis, L. G., Dibner, M. D. and Battey, J. F. (1986) *Basic Methods in Molecular Biology,* pp123. Elsevier

Efstratiadis, A., Kafatos, F. C. and Maniatis, T. (1977) *Cell,* 10: 571

Feinberg, A. P. and Vogelstein, B. (1983) *Anal. Biochem.*, 132: 6

Feinberg, A. P. and Vogelstein, B. (1984) *Anal. Biochem.*, 137: 266

Fieke, C. A. and Jencks, W. P. (1986) *J. Biol. Chem,* 261: 7603

Frischauf, A. M. , Lehrach, H., Poustka, A. and Murray, N. (1983) *J. Mol. Biol.*, 170:827

Frischauf, A. M. , Murray, N. and Lehrach, H. (1987) *Meth. Enzymol.*, 153: 103 Academic Press

Firtel, R. A. (1981) *Cell,* 24: 6

Fukao, T., Kamijo, K., Osumi, T., Fujiki, Y., Yamaguchi, S., Orii, T. and Hashimoto, T. (1989) *J. Biochem.*, 106: 197

Garoff, H. and Ansorge, W. (1981) *Anal. Biochem.*, 115: 450

Gibson, J., Upper, C. D. and Gunsalus (1976) *J. Biol. Chem*, 242: 2474

Goldsmith, M. R. and Kafatos, F. C. (1984) *Ann. Rev. Genet.*, 18: 443

Grossberger, D. (1987) *Nucl. Acid. Res.*, 15: 6737

Gubler, U. (1988) *Nucl. Acid. Res.*, 16: 2726

Gubler, U. and Hoffman, B. J. (1983) *Gene*, 25: 263

Haas, M. J. and Fleming, D. J. (1986) *Nucl. Acid. Res.*, 14: 3976

Haas, M. J. and fleming, D. J. (1988) *Anal. Biochem.*, 168: 239

Hastings, K., Bucher, E. and Emetson, Jr. C. (1985) *J. Biol. Chem.* 260: 13699

Helfman, D., Cheley, S., Kuismanen, E., Finn, L. and Yamawaki-Kataoka, Y. (1986) *Mol. Cell. Biol.*, 6: 3582

Helms,C., Dutchik, J. E. and Olson, M. V. (1987) *Meth. Enzymol.*, 153: 69 Academic Press

Henning, W. D., Upton, C., McFadden, G, Majumdar, R. and Bridger, W. A. (1988) *Proc. Natl. Acad. Sci. USA* 85: 1432

Hentschel, C. C. and Birnstiel, M. L. (1981) *Cell*, 25: 301

Hersh, L. B. and Jencks, W. P. (1967a) *J. Biol. Chem*, 242: 3468

Hersh, L. B. and Jencks, W. P. (1967b) *J. Biol. Chem.*, 242: 3481

Hong, G. H. (1987) *Method. Enzymol.*, 155: 93

Hood, L., Kornenberg, M. and Hunkapiller, T. (1985) *Cell*, 40: 225

Howard, J. B., Zieske, L., Clarkson, J. and Rathe, L. (1986) *J. Biol. Chem.*, 261: 60

Hunkapiller, M. , Lujan, E., Ostrander, F. and Hood, L. (1983) *Method. Enzymol.*, 91: 227

Hunkapiller, T. and Hood, L. (1986) *Nature*, 323: 15

Huynh, T. V., Young, R. A. and Davis, R. W. (1985) *DNA cloning: a practical approach* (Glover, D. M. , ed) 1: 49 IRL Press

Jackson, I. J. (1991) *Nucl. Acid. Res.*, 19: 3795

Jacobson, A. (1987) *Meth. Enzymol.*, (Berger, S. L. and Kimmel, A. R., ed) 152: 257 Academic press

Jeffreys, A. J., Wilson, V., Neumann, R. and Keyte, J. (1988) *Nucl. Acid. Res.*, 16: 10953

Jencks, W. P. (1973) *The Enzymes*, (Boyer, P. D., ed) 3rd ed. IX pp483 Academic Press

Jenkins, T. M. and Weitzman, P. D. J. (1986) *FEBS Lett.*, 205: 215

Karlik, C. and Fyrberg, E. (1986) *Mol. Cell. Biol.*, 6: 1965

Kimmel, A. R. and Berger, S. L. (1987) *Meth. Enzymol.*, (Berger, S. L. and Kimmel, A. R., ed) 152: 315 Academic Press

Kincaid, R. L. and Nightingale, M. S. (1988) *BioTech.*, 6: 42

Laemmli, U. K. (1970) *Nature*, 227: 680

Langer, P. R., Waldrop, A. A. and Ward, D. C. (1981) *Proc. Natl. Acad. Sci. USA*, 78: 6633

Lapeyre, B. and Amalric, F. (1985) *Gene*, 37: 215

Lathe, R. (1985) *J. Mol. Biol.*, 183: 1

Leary, J. J., Brigati, D. J. and Ward, D. C. (1983) *Proc. Natl. Acad. Sci. USA*, 80: 4045

Lewin, B. (1987) *Gene*, 3rd, John Wiley & Sons

MacDonald, R., Swift, G., Przybylo, A. and Chirgwin, J. (1987) *Meth. Enzymol.*, (Berger, S. and Kimmel, A. R., ed) 152: 129 Academic Press

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: a laboratory manual*, CSH Laboratory Press

Marusyk, R. (1980) *Anal. Biochem.*, 105: 403

Maruyama, T., Gojobori, T., Aota, S. and Ilemura, T. () Nucl. Acid. Res., 14 suppl.: r151

Matsudaira, P. (1987) *J. Biol. Chem.*, 262: 10035

McClellan, J. A. and Ottaway, J. H. (1980) *Comp. Biochem. Physiol.*, 67B: 679

McMinn, C. L. and Ottaway, J. H. (1979) *Biochem. Soc. Trans.*, 7: 411

Medford, R., Nguyen, H. T., Destree, A., Summers, E. and Nadal-Ginard, B. (1984) *Cell*, 38: 409

Messing, J. (1983) *Methods Enzymol.*, 101, part C : 20

Moffet, F. J., Wang, T. and Bridger, W. A. (1972) *J. Biol. Chem.*, 247: 8139

Moore, S. A. and Jencks, W. P. (1982) *J. Biol. Chem.*, 257: 10893

Moyer, R. W., Ramaley, R. F., Butler, L. C. and Boyer, P. D. (1967) *J. Biol. Chem.*, 242: 4299

Murakami, K., Mitchell, T. and Nishimura, J. S. (1972) *J. Biol. Chem.*, 247: 6247

Murakami, K. and Nishimura, J. S. (1974) *Biochim. Biophys. Acta*, 336: 252

Nabeshima, Y., Fujii-Kuriyama, Y., Muramatsu, M. and Ogata, K. (1984) *Nature*, 308: 333

Nasadini, R., Avogaro, A., Doria, A., Fioretto, P., Trevisan, R. and Morocutti, A. (1989) *Diabetes/Metabolism Rev.*, 5: 299

Nicholls, D., Sundaram, T., Atkinson, T. and Minton, N. (1988) *Nucl. Acid. Res.*, 16: 9858

Nishimura, J. S. (1986) *Adv. Enzy.*, 58: 141

Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) *Genetics*, 120: 621

O'Connor-McCourt, M. D., and Bridger, W. A. (1985) *Can. J. Biochem. Cell Biol.*, 63: 57

Okayama, H. and Berg, P. (1982) *Mol. Cell. Biol.*, 2: 161

Old, R. W. and Woodland, H. R. (1984) *Cell*, 38: 624

Olsson, A., Moks, T., Uhien, M. and Gaal, A. (1984) *J. Biochem. Biophys. Method.*, 10: 83

Ottaway, J. H., McClellan, J. A. and Saunderson, C. L. (1981) *Int. J. Chem.*, 13: 401

Periasamy, M. , Strehler, E. E., Garfinkel, L. I., Gubits, R. M. , Ruiz-Opazo, N. and Nadal-Ginard, B. (1984) *J. Biol. Chem.*, 259: 13595

Pickart, C. M. and Jencks, W.P. (1979) *J. Biol. Chem.*, 254: 9120

Porpaczy, Z., Sumegi, B. and Alkonyi, I. (1983) *Biochim. Biophys. Acta*, 749: 172

Proudfoot, N. (1986) *Nature*, 321: 730

Radding, C. M. , Josse, J. and Kornberg (1962) *J. Biol. Chem.*, 237: 2869

Reed, S., Hadwiger, J., Richardson, H. and Wittenberg, C. (1989) *J. Cell. Sci. Suppl*, 12: 29

Robert, B., Daubas, P., Akimenko, M-A., Cohen, A., Garnerr, I., Guenet, J-L. and Buckingham, M. (1984) *Cell*, 39: 129

Robinson, Jr. J. B. and Srere, P. A. (1985) *J. Biol. Chem.*, 260: 10800

Rougeon, F. and Mach, B. (1976) *Proc. Natl. Acad. Sci. USA*, 73: 3418

Ruiz-Opazo, N. and Nadal-Ginard, B. (1987) *J. Biol. Chem.*, 262: 4755

Ruiz-Opazo, N., Weinberger, J. and Nadal-Ginard, B. (1985) *Nature*, 315: 67

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science*, 230: 487

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: a laboratory manual*, 2nd ed. CSH Laboratory Press

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*, 74: 5463

Sharp, J. A. and Edwards, M. R. (1978) *Biochem. J.*, 173: 759

Sharp, P. (1981) *Cell*, 23: 643

Silver, J. and Keerikatte (1989) *J. Cell. Biochem.* (suppl.) 13E: 306

Smith, D. E. and Fisher, P. A. (1984) *J. Cell Biol.*, 99: 20

Solomon, F. and Jencks, W. P. (1969) *J. Biol. Chem.*, 244: 1079

Srivasava, D. K. and Brenhard, S. A. (1986) *Science*, 234: 1081

Strehler, E. E., Periasamy, M. , Strehler-Page, M-A. and Nadal-Ginard, B. (1985) *Mol. Cell. Biol.*, 5: 3169

Stryer, L. (1988) *Biochemistry*, 3rd ed. pp373 Freeman

Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K. and Wallace, R. B. (1981) *Developmental Biology Using Purified Genes* (Brown, D. and Fox, C. F.,ed.) pp683 Academic Press

Tabor, S. and Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA*, 84: 4767

Tabor, S. and Richardson, C. C. (1989) *J. Biol. Chem.*, 264: 6447

Triglia, T., Peterson, M. G. and Kemp, D. J. (1988) *Nucl. Acid. Res.*, 16: 8186

Vogel, H. J. and Bridger, W. A. (1982) *J. Biol. Chem.*, 257: 4834

Vogel, H. J. and Bridger, W. A.(1983) *Biochem. Soc. Trans.*, 11: 315

von Heijne, G. (1986) *EMBO J.*, 5: 1335

Wang, T., Jurasek, L. and Bridger, W. A. (1972) *Biochem.*, 11: 2067

Weitzman, P. D. J. (1987) *Biochem. Soc. Symp.*, 54: 33

Weitzman, P. D. J. and Jaskowskan-Hodges, H. (1982) *FEBS Lett.*, 143: 237

Weitzman, P. D. J., Jenkins, T., Else, A. J. and Holt, R. A. (1986) *FEBS Lett.*, 199: 57

White, H. and Jencks, W. P. (1976a) *J. Biol. Chem.*, 251: 1688

White, H. and Jencks, W. P. (1976b) *J. Biol. Chem.*, 251: 1708

White, H., Solomon, F. and Jencks, W. P. (1976) *J. Biol. Chem.*, 251: 1700

White, J. A., Stott, R. A. W. and Matthews, J. A. (1988) *Nucl. Acid. Res.*, 16: 2727

Wieczorek, D., Smith, C. and Nadal-Ginard, B. (1988) *Mol. Cell. Biol.*, 8: 679

Williamson, J. R. and Krebs, H. A. (1961) *Biochem. J.*, 80: 540

Wolodko, W. T., Brownie, E. R., O'Conor, M. D., and Bridger, W. A. (1983) J. Biol. Chem., 258: 14116

Wolodko, W. T., James, M. N. G. and Bridger, W. A. (1984) *J. Biol. Chem.*, 259: 5316

Wolodko, W. T., Kay, C. M. and Bridger, W. A. (1986) *Biochem.*, 25: 5420

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, 33: 103

APPENDIX I

Publications arising from my work:

I. Papers

1. LIN, T. & BRIDGER, W.A. Phosphorylation by cyclic AMP-dependent protein kinase does not affect the association of ATP-citrate lyase with isolated mitochondria. *FEBS Letters*, 224, 322-324 (1987).

2. LIN, T. & BRIDGER, W.A. Sequence of a cDNA clone encoding pig heart mitochondria CoA transferase. *J. Biol. Chem.* 267, 975-978 (1992).

3. LIN, T., RYAN, D.G. & BRIDGER, W. A. Alternative splicing produces variants of the mRNA encoding the α subunit of pig heart succinyl-CoA synthetase. *Mol. Cell. Biol.*, in preparation.

II. Abstracts

1. LIN, T., MAJUMDA, R., RYAN, D & BRIDGER, W.A. (1989) Import of succinyl coenzyme A synthetase into mammalian mitochondria, and demonstration of isoforms produced by alternative splicing mechanisms. Protein engineering '89 Second Internatioanl Conference, Japan.

2. LIN, T., RYAN, D.G., HENNING, W.D. & BRIDGER, W.A. (1990) Isoforms of succinyl-CoA synthetase can be generated by alternative splicing of the RNA transcript of its α subunit. Proc. Can. Fed. Biol. Sci., Abstract 194.

3. BAILEY, D., LIN, T. & BRIDGER, W.A. (1990) Cloning of the β subunit of succinyl-CoA synthetase from a pig heart cDNA library.

Abstracts, American Society for Cell Biology 30th Annual Meeting. *J. Cell Biol.* 5(2) p. 453a.

APPENDIX II

Some related research carried out by the author during the project:

1.Interactions of SCS and CoA transferase

According to the "metabolon hypothesis" (Robinson and Srere, 1985), one would predicted possible physical association between SCS and CoA transferase because the activities of two enzymes shared a common metabolic pathway. To explore this idea, we used three approaches that might detect specific interactions between these two enzymes.

a. Cross-linking experiments. Equal amounts of SCS and CoA transferase were incubated with cross-linkers (dimethyl suberimidate, glutaraldehyde, etc.). The cross-linked products were analyzed on SDS PAGE. The results indicated that mixtrues of SCS and CoA transferase produced no new products in cross-linking reactions compared to SCS and CoA transferase samples that were cross-linked separately.

b. Active enzyme ultracentrifugation. The experiments were done essentially according to Wolodko, *et al.* (1986), to investigate possible high molecular weight complexes associated with SCS activity. The results indicated that there was no larger SCS species when this enzyme was co-centrifuged with CoA transferase, compared to the sedimentation pattern when it was centrifuged alone.

c. Kinetic studies. It was demonstrated that the K_m's of CoA transferase were unaffected by the presence of SCS.

Conclusion: there were no interactions between SCS and CoA transferase that were detected by the techniques used.

2. Hybrid SCS

The possible function associated with the 108-cassette could be investigated by engineering the sequence into the clones of *E.coli* SCS which had been over-expressed. The author made two constructs for this purpose. One of the constructs replaced the sequence coding for the α subunit of *F.coli* SCS with that of pig heart SCS containing the PS108 cassette. The other hybrid construct, to be used as a control, had the sequence encoding the pig heart α subunit with the PS57 cassette in the *E.coli* counterpart. However, it was found that both of the hybrid α subunits were not expressed, despite the fact that the preceding β gene was expressed normally.

3. Cloning the α subunits of SCS from different species by PCR

When the cDNA sequences of the α subunits of pig heart, rat liver and *E.coli* were all available and compared, it was found that there were two stretches of sequence (one stretch near the 5' end and the other near the 3' end of the coding sequences) that were similar, some even in all 3 positions of the codons. A strategy was developed to design two degenerate oligonucleotides which could best represent the sequences with manageable species to be used in PCR experiments to clone α subunit genes from different species. The goal of this approach was to obtain with relative ease the sequences of the α subunits of SCS from a variety of species, from lower prokaryotes to higher eukaryostes and to use the information to better understand structure and function relationships. Using yeast DNA as template, a fragment was amplified with the expected size. Restriction analysis indicated that this DNA fragment was not the same as that derived from rat liver, pig heart or *E.coli* clones. There are also additional clones of the α subunits of SCS (yeast, *thermoaquatics*) available. Better primers can be derived and this approach may still be profitable.